

Kant and Religion

This masterful work on Kant's *Religion within the Boundaries of Mere Reason* explores Kant's treatment of the idea of God, his views concerning evil, and the moral grounds for faith in God. *Kant and Religion* works to deepen our understanding of religion's place and meaning within the history of human culture, touching on Kant's philosophical stance regarding theoretical, moral, political, and religious matters. Wood's breadth of knowledge of Kant's corpus, philosophical sharpness, and depth of reflection sheds light not only on Kant but also on the fate of religion and its relation to philosophy in the modern world.

Allen W. Wood is Ruth Norman Halls Professor of Philosophy at Indiana University.

Cambridge Studies in Religion, Philosophy, and Society

Series Editors

Paul K. Moser, *Loyola University, Chicago*

Chad Meister, *Bethel College, Indiana*

This is a series of interdisciplinary texts devoted to major-level courses in religion, philosophy, and related fields. It includes original, current, and wide-spanning contributions by leading scholars from various disciplines that (a) focus on the central academic topics in religion and philosophy, (b) are seminal and up-to-date regarding recent developments in scholarship on the various key topics, and (c) incorporate, with needed precision and depth, the major differing perspectives and backgrounds – the central voices on the major religions and the religious, philosophical, and sociological viewpoints that cover the intellectual landscape today. Cambridge Studies in Religion, Philosophy, and Society is a direct response to this recent and widespread interest and need.

Recent Books in the Series

Roger Trigg *Religious Diversity: Philosophical and Political Dimensions*

John Cottingham *Philosophy of Religion: Towards a More Humane Approach*

William J. Wainwright *Reason, Revelation, and Devotion: Inference and Argument in Religion*

Harry J. Gensler *Ethics and Religion*

Fraser Watts *Psychology, Religion, and Spirituality: Concepts and Applications*

Gordon Graham *Philosophy, Art, and Religion: Understanding Faith and Creativity*

Keith Ward *The Christian Idea of God: A Philosophical Foundation for Faith*

Timothy Samuel Shah and Jack Friedman *Homo Religiosus? Exploring the Roots of Religion and Religious Freedom in Human Experience*

Sylvia Walsh *Kierkegaard and Religion: Personality, Character, and Virtue*

Roger S. Gottlieb *Morality and the Environmental Crisis*

J. L. Schellenberg *Religion after Science: The Cultural Consequences of Religious Immaturity*

Clifford Williams *Religion and the Meaning of Life: An Existential Approach*

Kant and Religion

ALLEN W. WOOD

Indiana University

Stanford University



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE

UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781108422345

DOI: 10.1017/9781108381512

© Allen W. Wood 2020

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2020

Printed in the United Kingdom by TJ International Ltd., Padstow, Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

NAMES: Wood, Allen W., author.

TITLE: Kant and religion / Allen W. Wood.

DESCRIPTION: 1. | New York : Cambridge University Press, 2020. | Series: Cambridge studies in religion, philosophy, and society | Includes bibliographical references and index.

IDENTIFIERS: LCCN 2019051968 (print) | LCCN 2019051969 (ebook) |

ISBN 9781108422345 (hardback) | ISBN 9781108381512 (ebook)

SUBJECTS: LCSH: Kant, Immanuel, 1724-1804 – Religion. | Metaphysics. | Immortality (Philosophy) | Soul – Philosophy. | God – Controversal literature.

CLASSIFICATION: LCC B2799.R4 W65 2020 (print) | LCC B2799.R4 (ebook) | DDC 210.92-dc23

LC record available at <https://lcn.loc.gov/2019051968>

LC ebook record available at <https://lcn.loc.gov/2019051969>

ISBN 978-1-108-42234-5 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

To the memory of my brother

Roger Earl Wood

(1947–2008)

Contents

<i>Citations</i>	<i>page</i> x
<i>Preface</i>	xiii
<i>Note on the Cover Image</i>	xxii
1 Religion and Reason	I
1.1 What Is “Religion”?	2
1.2 Religion as Essentially Symbolic	4
1.3 Religion in Kant’s Life	9
1.4 The Experimental Testing of a Hypothesis	11
1.5 The First Experiment	14
1.6 The Second Experiment	16
1.7 The Publication of Kant’s <i>Religion within the Boundaries of Mere Reason</i>	20
1.8 Wöllner’s Charges and Kant’s Defense	22
2 Moral Faith in God	27
2.1 Philosophical Arguments for Belief in God	27
2.2 Practical Grounds of Assent	29
2.3 What Is at Stake in Kant’s Practical Arguments	34
2.4 The Highest Good as a Moral End	37
2.5 The Moral Argument in the <i>Critique of Pure Reason</i>	39
2.6 The Moral Argument in the <i>Critique of Practical Reason</i>	45
2.7 The Moral Argument in the <i>Critique of the Power of Judgment</i>	49
2.8 The Kind of Assent Justified by the Moral Argument	51
2.9 Is Practical Faith Antievidentialist?	56
3 The Radical Evil in Human Nature	61
3.1 Original Sin and Radical Evil	61
3.2 Practical Freedom	63
3.3 The Concept of Moral Evil	66

3.4	The Propensity to Evil	70
3.5	Kant's "Rigorism"	73
3.6	Walpole's Thesis	75
3.7	Evil and Unsociable Sociability	76
3.8	The Passions	79
3.9	The Three Degrees of Evil	82
3.10	Use of the Thesis of Radical Evil	84
4	The Change of Heart	89
4.1	The Purity of the Moral Incentive	90
4.2	Goodness of Heart	91
4.3	Acting from an Incentive	94
4.4	Acting from Duty in the <i>Groundwork</i>	98
4.5	Legality and Morality	101
4.6	The Narrow Duty to Act from Duty	104
4.7	The Wide Duty to Act from Duty	107
4.8	Not a Datable Event in Time	110
5	The Son of God	115
5.1	The Moral Ideal	115
5.2	Symbols and Analogy	117
5.3	Symbols and Religion	120
5.4	The Son of God as a Symbol	124
5.5	Faith in the <i>Reality</i> of the Son of God	126
5.6	The Three Difficulties	127
5.7	God's Decree of Grace	133
5.8	What Is Forgiveness?	135
6	Grace and Salvation	139
6.1	Justifying Grace and God's Forgiveness	140
6.2	Vicarious Atonement	143
6.3	Jesus Christ as the Son of God	148
6.4	Salvation by Grace Alone	151
6.5	Moral Presuppositions of Grace	152
6.6	God's Causality and the "Inconsistent Quartet"	155
6.7	Augustinianism about Grace	161
7	The Ethical Community and the Church	164
7.1	The Ethical State of Nature	164
7.2	The Ethical Community	166
7.3	Ecclesiastical Faith	168
7.4	The Need for Revealed Religion	170
7.5	Sublimity and Revealed Religion	172
7.6	The Interpretation of Scripture	174
7.7	Faith and Works: The Antinomy	178
7.8	The Coming of the Kingdom of God	182

8	Freedom of Conscience	185
8.1	What Is Service of God? What Is “Counterfeit” Service?	187
8.2	Mendelssohn and Kant	190
8.3	Mendelssohn on Natural Right, Church, and State	192
8.4	Mendelssohn’s “New and Harsh” Doctrines	195
8.5	Freedom of Conscience and Evidentialism	197
8.6	Mendelssohn’s Enlightened Judaism	200
8.7	Where Mendelssohn and Kant Differ	201
8.8	Kant on Religious Conscience	204
	Concluding Remarks	210
	<i>Bibliography</i>	216
	<i>Index</i>	225

Citations

Writings of Kant will be cited according to the following abbreviations:

- Ak *Immanuel Kants Schriften*. Ausgabe der königlich preussischen Akademie der Wissenschaften (Berlin: W. de Gruyter, 1902–). Unless otherwise footnoted, writings of Immanuel Kant will be cited by volume:page number in this edition.
- Ca *Cambridge Edition of the Writings of Immanuel Kant* (New York: Cambridge University Press, 1992–2016). This edition provides marginal Ak volume:page citations. Specific works will be cited using the following system of abbreviations. (Works not abbreviated below will be cited simply as Ak volume:page.)
- Anth *Anthropologie in pragmatischer Hinsicht* (1798), Ak 7 *Anthropology from a Pragmatic Point of View*, Ca Anthropology, History, and Education
- ED *Das Ende Aller Dinge*, Ak 8 *The End of All Things*, Ca Writings on Religion and Natural Theology
- EF *Zum ewigen Frieden: Ein philosophischer Entwurf* (1795), Ak 8 *Toward Perpetual Peace: A Philosophical Project*, Ca Practical Philosophy
- G *Grundlegung zur Metaphysik der Sitten* (1785), Ak 4 *Groundwork of the Metaphysics of Morals*, Ca Practical Philosophy
- I *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht* (1784), Ak 8 *Idea toward a Universal History with a Cosmopolitan Aim*, Ca Anthropology, History, and Education
- KrV *Kritik der reinen Vernunft* (1781, 1787). Cited by A/B pagination. *Critique of Pure Reason*, Ca Critique of Pure Reason
- KpV *Kritik der praktischen Vernunft* (1788), Ak 5 *Critique of Practical Reason*, Ca Practical Philosophy
- KU *Kritik der Urteilskraft* (1790), Ak 5 *Critique of the Power of Judgment*, Ca Critique of the Power of Judgment

- MA *Mutmaßlicher Anfang der Menschengeschichte* (1786), Ak 8 *Conjectural Beginning of Human History*, Ca Anthropology, History, and Education
- MS *Metaphysik der Sitten* (1797–1798), Ak 6 *Metaphysics of Morals*, Ca Practical Philosophy
- O *Was heißt: Sich im Denken orientieren?* (1786), Ak 8 *What Does It Mean to Orient Oneself in Thinking?* Ca Religion and Rational Theology
- P *Prolegomena zu einer jeden künftigen Metaphysik* (1783), Ak 4 *Prolegomena to Any Future Metaphysics*, Ca Theoretical Philosophy after 1781
- R *Religion innerhalb der Grenzen der bloßen Vernunft* (1793–1794), Ak 6 *Religion within the Boundaries of Mere Reason*, Ca Religion and Rational Theology
- SF *Streit der Fakultäten* (1798), Ak 7 *Conflict of the Faculties*, Ca Religion and Rational Theology
- TP *Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis* (1793), Ak 8 *On the Common Saying: That May Be Correct in Theory But It Is of No Use in Practice*, Ca Practical Philosophy
- VA *Vorlesungen über Anthropologie*, VA 25 *Lectures on Anthropology*, Ca Lectures on Anthropology
- VE *Vorlesungen über Ethik*, Ak 27, 29 *Lectures on Ethics*, Ca Lectures on Ethics, by transcriber: Collins, Mrongovius, Vigilantius
- VL *Vorlesungen über Logik*, Ak 9, 24 *Lectures on Logic*, Ca Lectures on Logic
- VM *Vorlesungen über Metaphysik*, Ak 27 *Lectures on Metaphysics*, Ca Lectures on Metaphysics
- VP [*Vorlesungen über*] *Pädagogik*, Ak 9 *Lectures on Pedagogy*, Ca Anthropology, History, and Education
- VpR *Vorlesungen über die philosophische Religionslehre*, Ak 28 *Lectures on the Philosophical Doctrine of Religion*, Ca Religion and Rational Theology
- VRL *Über ein vermeintes Recht aus Menschenliebe zu lügen*, Ak 8 *On a Supposed Right to Lie from Philanthropy*, Ca Practical Philosophy
- WA *Beantwortung der Frage: Was ist Aufklärung?* (1784), Ak 8 *An Answer to the Question: What Is Enlightenment?* Ca Practical Philosophy

Preface

The topic of this book is Kant's philosophical project in *Religion within the Boundaries of Mere Reason*. Its main focus is Kant's philosophical engagement with religious faith and Christian theology as he knew them. But Kant's project in the *Religion* will take us into other aspects of Kant's metaphysics and moral philosophy. We will encounter the unknowability of the supersensible and the roles of analogy, symbolism, and aesthetic experience in our thinking about it. The theme of religion will take us into Kant's theory of rational assent and the importance of enlightenment as thinking for oneself. We will deal with important topics in Kant's moral philosophy and moral psychology, such as moral evil, the distinctiveness of the moral incentive, and freedom of conscience. We will also have to appreciate some themes in Kant's ethics that have been underemphasized or even badly distorted, such as the connection of the moral incentive to emotions and goodness of heart, the human experience of guilt and the challenge of self-acceptance, and the vital importance of community in Kant's ethics and of the hope for moral progress of the human species in history. The wide range of these topics indicates the depth and centrality of Kant's thinking about religion for his entire philosophy.

In addition to Kant's project in the *Religion* in relation to its own time, however, the book will inevitably have another focus. Throughout we will be addressing, at least implicitly, also what Kant's treatment of religion might mean for us. That includes what Kant's philosophy, as regards its religious import, has meant for our culture in the last two centuries. The present book is not the place to review that history. So I choose to leave much of this second, inevitable focus implicit. I will be challenging the idea that Kant's philosophy of religion was meant to advance the cause of "secularism" in the modern world, or to rid modern culture of religion and replace it with a nonreligious rational morality. I will certainly not be looking at Kant – as Gordon Michalson puts it – as "a way station between Luther and Marx" (Michalson, 1999, p. 27). I shamelessly profess to think that this would be a route the human race has

not traveled but should have traveled. But you won't find me directly defending that thought here. A responsible book on Kant's philosophy of religion cannot address such issues except indirectly.

This is not a biographical study of Kant. In Chapter 1, I do discuss Kant's religious background and his dealings with the Prussian government over the *Religion's* publication, but only because these matters shed light on Kant's philosophical project. Criticism or defense of Kant's personality, his attitudes on sex and race, or even his personal religious attitudes are not part of my project. I won't be trying to decide whether Kant should be criticized for being "too sympathetic to religion" or "not sympathetic enough," since I will leave it to my readers to decide, each for him- or herself, how sympathetic to religion someone ought to be. Those who approach Kant (or any other great philosopher) from the standpoint of what they like to call the "hermeneutics of suspicion" represent only the forces of anti-intellectualism and hate. We live in an age of religious hate, which includes antireligious hate. I don't doubt that some people reading Kant's project in the *Religion*, or what I have to say about it, will find ways of interpreting and self-righteously rejecting it from some hate-driven standpoint of "moral clarity." There is nothing I can do to prevent them, except to protest in advance and express the wish that they would not do it. For those interested in what I have to say about Kant as a human being, see Wood, 2005, pp. 1–23 and Wood, 2008, pp. 6–12.

Nor do I intend to survey recent scholarship on Kant's philosophy of religion, which is surprisingly large and still growing. I did not in the least anticipate that Kant's philosophy of religion would become such a focus of scholarly interest at the time I wrote my first book about Kant's religious thought half a century ago. Though some of this literature is very good, and I have learned much from it, it is also a sadly true generalization that whenever literature on a subject explodes in this manner, a great deal of it is inevitably not of high quality. I refer to the literature only occasionally, when I think mention of it sheds light on what I want to say. Surveys of the literature prior to the last ten years have been done by Palmquist (2000); Firestone and Palmquist (2000); Firestone and Jacobs (2008); and Firestone (2009, chap. 1).

There has been a stream of recent scholarship favorable to Kant and at the same time to traditional Christianity, represented by such people as Chris Firestone, Nathan Jacobs, Stephen R. Palmquist, and Andrew Chignell. I agree with some of this scholarship and disagree with some. I regard it as worse than unfortunate, however, that its generally sympathetic attitude toward Kant's *Religion* has been so atypical of what people think about Kant and religion. Secular philosophers have often seen Kant's sympathy with religious belief, and Christianity in particular, as a reason for keeping their distance from Kant. They suspect Kant's categorical imperative of reintroducing into ethics the supernatural moral tyrant of traditional religion without openly admitting it. Or else they react as Goethe did when he accused Kant of "wantonly tainting [his philosophy] with the shameful stain of radical

evil in order that even Christians might be attracted to kiss its hem” (Goethe, 1988, 2:166).

Traditional religious believers have been even less kind to Kant on religion. They usually view Kant as an enemy. Thus Gordon Michalson has written that Kant’s “efforts to ameliorate the theologically destructive effects of the *Critique of Pure Reason* implicitly make things worse for Christian faith, not better” (Michalson, 1999, p. 5). Even more bluntly hostile is Matthew Alun Ray, whose book’s title leaves nothing to the imagination: *Subjectivity and Irreligion: Atheism and Agnosticism*. Ray concludes that Kant’s philosophy of religion leaves us with only “Königsbergian nihilism” (Ray, 2003, p. 26). Treatments of Kant on the Christian religion by more scholarly Christian writers such as Nicholas Wolterstorff, John Hare, and Christopher Insole stress what they see as “conundrums,” gaps, inconsistencies, and failures in Kant’s project. Even one recent writer who professes to be defending Kant against such criticisms apparently cannot take Kant’s basic project seriously in the terms Kant states it. He concludes that Kant’s own declaration in his response to the Prussian authorities that he was not disparaging Christianity was disingenuous, “not entirely innocent” of “guile” and “prevarication” (Pasternack, 2014, p. 12).

The views people attribute to Kant on religion often tell you far more about those people than they do about Kant. In this presentation of Kant on religion, I am trying to let Kant speak, with my only additions being the intellectual sympathy and at the same time the critical perspective that any philosopher should always bring to the study of another philosopher. But perhaps readers of this book will think it is true of me, too, that my views on religion color and distort. Therefore, most of this Preface will be used in an attempt to get those issues out of the way by putting my cards on the table, so that my readers can decide for themselves whether my views are distorting my intended aims. Thus I offer the following five or so confessional pages in the interest of full disclosure, and not because I think other people ought to be particularly interested in my religious autobiography for its own sake.

I was raised a Christian – more specifically, an Episcopalian (or Anglican). I was, I think, a sincere religious believer until about the age of twenty. As a Christian adolescent, I saw myself as a “middle-church” Anglican, always at some distance both from Roman Catholicism and from all “fundamentalist” or “evangelical” sects. The earliest writers who influenced me about religion were Thoreau and Tolstoy. During this time I thought all forms of secularism shallow and dehumanizing. I was also always a firm believer in individual religious conscience, and suspicious of authority in matters of religion – as well as of authority more generally. We should listen with respect to the thoughts of anyone whose competence and arguments, on their merits, justify that respect. Then we should draw our own conclusions. In no area of life have I ever thought it morally acceptable for any human adult to defer submissively to the thoughts of others because these others hold some institutional position, whether priest, professor, or politician.

I recently read an op-ed piece arguing that Americans choose their politics before they choose their religion. This was presented as a recent development, but I think it has been true for a long time. In the US Civil War, both the proslavery and antislavery causes were motivated by what they thought were Christian convictions. Perhaps this same connection between religious and political convictions has held for me, too, though in that case it included some significant political self-discovery. It was sincere religious conviction that led me to be a pacifist (a conscientious objector to military service) even before the Vietnam War made that stance fashionable. I even announced my refusal of military service so early that the FBI still had the resources to investigate me and put an official bureaucratic stamp on my religious sincerity. But student and occupational draft deferments always kept me one step ahead of either an induction order or a requirement to do substitute civilian service. I did not do as both my brother and brother-in-law did and do alternative service itself as evidence to myself of my conscientious antiwar convictions. Like Vice President Dick Cheney, I had “other priorities.” Unlike him, however, I have never wanted to send others to kill and die in my place.

In those days I considered myself politically very conservative. This self-conception lasted for a while even after people on both the right and the left persistently informed me that my moral and political views made me a “liberal” or even a “leftist.” The views they meant were: my pacifism, my sympathy with the civil rights movement, and my belief that any significant economic inequality and especially the very existence of poverty in an opulent society are outrageous injustices for which there can be no excuse. I believed that all these convictions must be consistent with political conservatism simply because they are absolutely required by the Gospels and by basic human decency. That last belief has not changed at all. But it took me several years during this same time to change my own view of where I belong on the political spectrum, which I did as I came to add Marxism and feminism to my strongly held political convictions. It then took me longer still to drift gradually away from my religious upbringing. Some might say I “lost my Christian faith.” If so, this was more like losing ten unwanted pounds, or losing the mugger (or cop car) pursuing me, than like losing my keys, my wallet, or my way. I remain a pacifist to this day, though the original religious basis for it is long gone.

In Chapter 2 §3, I try to explain in what sense Kant’s moral argument shows faith in something to be necessary for any decent and thinking person. I did not know it at the time, but my first book, *Kant’s Moral Religion*, which dealt sympathetically with this argument, was a stage in my leave-taking of my religious upbringing. Its young author still had a great deal to learn about Kant’s ethics, since he accepted many errors about it that were current then and are sadly still current now. I have told people that I no longer feel I am identical with the author of that book. That’s another reason why I am writing this one.

The book I have taught in my classes more than any not authored by Immanuel Kant is Dostoevsky's greatest novel, *The Brothers Karamazov*. Central to its philosophical and religious argument is Ivan Karamazov's never-written poem, *The Legend of the Grand Inquisitor*, with its image of the aged Inquisitor, who uses institutionalized Christianity to serve Satan rather than Christ. The idea that human attempts to appropriate Christianity through organized religion might be seen as mortally hostile to Christ has seemed compelling to me ever since I first read Karl Barth's *Epistle to the Romans* (1919) as an undergraduate.

Since the 1970s, it has been clear to me that as a political force in the United States, white evangelical Christianity is thoroughly pernicious. I am not forgetting African American evangelicals, "Faithful America," or the #exvangelicals. And I am sure there are (because I have known some of them) individual white evangelicals who do sincerely adhere to the values of the Gospels. So let's be clear: I am talking about white evangelical Christianity exclusively *as a political movement* when I say without hesitation or qualification that it has no redeeming virtues. It is driven by toxic delusions, the motivation of which is dark, twisted, based on fear and hatred, openly barbaric, openly tyrannical, valuing cruelty toward the vulnerable for its own sake, even rejecting all democratic traditions and the rule of law. Its violent inversion of the values of the Gospels is sick, monstrous, and nihilistic.

Sometimes people find themselves slipping into an evil or criminal enterprise, and as their activities progress, perhaps as they even achieve some success, they eventually find themselves accepting and even committing evils which they could not even have conceived possible when they began. White evangelical Christianity in the late twentieth and early twenty-first centuries exemplifies this kind of downward spiral. The cruelty and malice of their policies are thinly disguised (but only from themselves) by the outrageous hypocrisy which has long been sadly common among religious people of all denominations. Opposition to legal abortion is just about the only political cause that they could even pretend to represent to anyone as anything but wanton cruelty. But it is no exception. The pretense of moral decency, which they offer by saying they "care about the babies," is a transparent lie in view of what they are actually doing. They are claiming legal guardianship over part of another person's body and claiming it for the sole purpose of depriving that very person of her freedom. Then, after violating her most personal and most fundamental rights, they refuse to take any responsibility for what would actually become a "baby" (by being born). A more exquisitely depraved combination of callousness, cruelty, and hypocrisy would even be hard to imagine.

I think many of these people know they have allied themselves with evil. But they think (or at least sometimes *say* they think) that they believe God will use this evil (which they sometimes say they regret) to bring about good. Some have compared it to Persian Emperor Cyrus in 2 Chronicles 36. Cyrus was a pagan tyrant, who nevertheless ended the Babylonian Captivity and ordered the

rebuilding of the Temple in Jerusalem. There is at work here a familiar and famously corrupt principle: *the end justifies the means*. Or as St. Paul puts it: “Why not say as some slanderously say (as some slanderously report that we say): Let us do evil that good may come? whose damnation is just” (*Romans 3:6*).

What is corrupt here is not only the use of wrongful means to achieve what can be represented as good ends. For then there would still possibly be a question whether the desirability of the end might outweigh the undesirability of the wrongful means; and it is at least imaginable that it might. What is really corrupt about *the end justifies the means*, or *doing evil that good may come*, is instead a dynamic all too familiar in political movements. Devotion to what seemed a good end is transformed into merely a means to wrongfulness, which then, objectively regarded, becomes to any clear-eyed observer the true end of the movement. The allegedly noble intention with which the evildoers supposedly began has been corrupted at its core. This dynamic has certainly shown itself in obvious – and terrible – ways in movements and political regimes that have called themselves Marxist. There the dogma of a historical process with its own “objective” teleology, like the faith in divine providence among religious people, serves as a pretext for self-deception. Those who think they know “whose side God is on” applaud barbarism as what they represent as God’s chosen means to do his benevolent will.

The only way to make sure you are pursuing a good end is to pursue it only through actions you can regard as in themselves right and good. Choose only good ends and also good means to them. This of course makes many desirable ends more difficult to achieve, perhaps even impossible. But it is a hard fact that the good is never easy to achieve, and sometimes we really do lack the resources to achieve desirable ends by permissible means. I might point out that the thoughts just expressed could constitute the essence of political *conservatism*, to the extent that the word denotes anything admirable. In that respect, despite my sympathy with Marxism and with other radical hopes, I remain a conservative to this day.

The long-term influence of white evangelical Christianity as a political movement may turn out to be the very opposite of what it desired (or thought it desired). From talking to my students, I fear that it has turned an entire generation of decent young people away from Christianity and even away from all religion. I now have a hard time getting most students to take seriously any philosopher, such as Descartes, Kant, or Kierkegaard, who even uses the “G-word.” I understand their attitude, but as a teacher and scholar of the history of philosophy, it is my job to help them to see how good and intelligent people in the past, and even today, could be religious believers. Their difficulty in grasping this is understandable during a time when in popular culture Christianity has been identified with all the most despicable of ends.

Religion is, and it ought to be, about emotions even more than it is about intellectual convictions or even moral resolve. I hope that in this book we will see that this is Kant's view too. Consequently, the cultural and political influence of a corrupt version of Christianity in my time and nation could not avoid having a powerful emotional impact on me. I hope I can still appreciate intellectually much of what is true, good, and beautiful in Christian symbols; I make a principled effort to do so when I encounter these symbols in literature, art, or philosophy. I hope the reader will see this effort on display in this book as I try to present Kant's positive engagement with the Christian religion in a sympathetic light. Kant saw plenty that was wrong with the Christianity of his time, but he had no difficulty seeing the Christian religion as a positive moral force in the life of individuals. He even invested in it his deepest hopes for progress in human history. It takes considerable intellectual effort for me to imagine sharing his sentiments on those points, but I think the effort, and even the need for it, has helped me to understand Kant better. In this book I try to present sympathetically his favorable attitude toward Christianity. But the political exploitation of Christianity for the most evil causes has fatally poisoned Christianity for me as regards my own life. Perhaps the thing they can't seem to poison for me is great religious music. I would begin by naming Brahms's *Deutsches Requiem*, and then go on to works by Bach, Bloch, Britten, Fauré ... but that would take up too much space.

Although I do not claim to be a religious person myself, I do respect many religious people, and I humbly (though also skeptically) remain open to their opinion that their religion is the foundation of what I know to be good about them. I mention here my long-time friends Robert Merrihew Adams and the late Marilyn McCord Adams, who have placed Christian faith at the very center of their lives. My wife's Jewish grandmother became a Quaker and raised my wife as one. My wife used to attend Quaker Meeting regularly – especially during the twenty years she was affiliated with a Roman Catholic university. I have known many admirable Catholics and Quakers. I have also known a fair number of practicing Jews and at least a few Muslims, whose religious beliefs and practices were an important and positive part of their admirable lives. The person I proudly regard as *my very first student*, Janice Dean Willis, considers herself to be *both* a Tibetan Buddhist and a black Baptist. Her autobiography, *Dreaming Me: Black, Baptist, and Buddhist* (Willis, 2008), makes for rewarding reading by anyone.

I therefore resist the antireligious attitude adopted by most people with whom I associate in the academic and scientific community. A self-inflicted intellectual incapacity to appreciate the human potential of religion is a serious deficiency in the culture of most scientists and most academic humanists. Studying Kant or others who combine reason, science, and enlightenment with religious faith helps me, and should help others, toward a more positive view of religion.

I also have a long-standing philosophical commitment to what is sometimes called *evidentialism*. This is the moral conviction that, in the words of W. K. Clifford (its most famous historical exponent), “it is wrong, always everywhere and for anyone to believe anything upon insufficient evidence” (Clifford, 1999, p. 77). Or to adopt a formulation closer to that of David Hume: we should always *proportion our beliefs to the evidence* (Hume, 2007, p. 80). Two other historically important evidentialists (both of them religious believers) were John Locke and Moses Mendelssohn. The views of the latter will be discussed in Chapter 8.

“Evidence,” as I understand it in this context, includes those theoretical or epistemic grounds, whether rational, empirical, or transcendental. Kant’s “deduction” of freedom and the moral law, for instance, is a philosophical argument to the effect that we would fall into incoherence if we did not presuppose that we are free – that is, able to judge and act for reasons, and especially to be motivated to obey a moral law of which we may regard our own rational will as the author. Avoiding incoherence is a reason to believe. I do not regard assent based on such philosophical arguments as violating the evidentialist principle. What evidentialism excludes, and treats as a moral failing, are the following: beliefs motivated by wishful (or fearful) thinking, beliefs based on partisanship, and beliefs you manipulate yourself into holding because you think they support some flattering conception of yourself or because it benefits you to hold them. None of these is evidence or any reason to think that what you believe is true.

I would equally insist that it is impossible altogether to exclude such irrational and corrupting influences on our beliefs. It is even an aspect of our freedom to act on reasons that we are also free not to act on them. Reasons differ from causes in that causes always necessitate, while reasons never do. Beings who act on reasons are free to choose which reasons they think are good reasons, and therefore are free to choose to act on bad reasons or to refuse to act on reasons at all. For imperfectly rational beings like ourselves, the freedom to act rationally is always the freedom to act irrationally. Fichte would call the freedom to act for reasons or not “formal” freedom, while he would call the higher or fuller freedom to act for good reasons “material freedom.” Part of being an evidentialist, as I understand the position, is being constantly aware (as far as you are able) of the influences on you to act contrary to reason and doing your best to resist them, while at the same time being aware that human fallibility being what it is, you will never be able entirely to do so. When I turned away from religious faith, I did so partly because it often deliberately chooses to violate the evidentialist principle. Many interesting thinkers – Pascal, Kierkegaard, Dostoevsky, and James – were outspoken violators of the evidentialist principle. In Chapter 2, I attempt to show how Kantian moral faith and evidentialism, while in tension, are not ultimately irreconcilable: in fact, they are complementary virtues.

I here find myself writing about the same topic with which I began my career in the 1960s. I hope I can now say better what attracts me to this topic. Kant had

hopes about the future of politics and also about the future of religion. Living under a military despotism, and surrounded by similar authoritarian regimes, he believed that people would come to see that the only legitimate form of political constitution is a representative republic, constituted so as to protect the rights of individual human beings. Living at a time when he saw religion as still in the chains of “priestcraft,” Kant looked forward to a time when religion would be the chief vehicle for humanity’s collective devotion to the cause of enlightenment. Religious communities would unite human beings in service of reason, helping humanity to progress toward the ideal of a realm of ends, a “kingdom of God on earth.”

I think we can now see that Kant was historically right about politics: although even the flickering hope for the state he regarded as legitimate has now become extremely fragile, most educated people in the world today would still agree that a representative republic protecting the rights of all citizens and open to the equal participation of all is the political state we ought to try to create. About religion, however, Kant would seem to have been historically wrong. Neither institutionalized religion nor institutionalized reason in the form of science and learned culture have changed in the ways Kant hoped they would. Yet when a philosopher seems to have been wrong about the course of the world, it’s not necessarily the philosopher who has gone wrong. The philosopher may have been right, and it may be that it is the course of the world that went wrong. As you will see if you read this book, this is my basic thought about Kant and religion. But another aim of the book is to correct the basic error of trying to decide whether Kant accepts or rejects Christianity, or this or that Christian doctrine. But that question, however we may answer it, gets Kant’s project entirely wrong. His aim regarding revealed Christianity is not simply to accept or reject it but ultimately to accept it while interpreting it critically.

Kant’s Moral Religion was written under the influence of my undergraduate religion teacher at Reed College, Daniel L. Deegan. Deegan died very young from cancer, before my dissertation was even completed, so that my first book was dedicated to his memory. One of his classmates at Yale Divinity School in the 1950s was my friend Van Harvey, who has given me much-valued comments on parts of this book. I have also received helpful comments, advice, and questions from Andrew Chignell, Karl Ameriks, Alyssa Bernstein, Sandra Shapshay, Samuel Kahn, Arthur Ripstein, Michael Morgan, Olga Lenczewska, Kimberly Brewer, Desmond Hogan, and the following students at Princeton University: Brendan Kolb, Carrie Pritt, Enoch Kuo, Rochhuahthanga Jongte, Cole Diehl, Asad Zafar Haider, Alejandro Naranjo Sandoval, Haley Brennan, and Kevin Zhou.

Cover Image

Cover image: *Königsberg, Löbenicht*, 2009, signed hand-printed woodcut by Sebastian Harwardt.

“At six o’clock Kant sat down at his desk, which was a simple, ordinary table, and read until dusk. At this time, which he found so conducive to thought, he would meditate on what he was reading, if it was worth that, or he would sketch out what he intended to say in his lectures the following day, or he would work on something meant for publication. Then, whether it was winter or summer, he would sit by the stove from which he could see through his window the Löbenicht steeple. He would contemplate it while meditating, or rather, let us say, he would rest his eyes on it. He could not emphasize enough now good for his eyes this was – how suitable the distance of the object was for this purpose. His daily gaze in the twilight accustomed his eyes to it. But eventually some poplars in his neighbor’s garden grew to such a height that they hid the steeple, which left Kant unsettled, and disturbed his meditation; so he asked that the poplars be pruned. Fortunately, the owner of the garden was a generous person who loved and respected Kant, and for his sake he sacrificed some of his poplar boughs, making the steeple visible once more, so that Kant could once again meditate undisturbed.”

E. A. C. Wasianski, *Immanuel Kant in seinen letzten Lebensjahren* (Königsberg: Nicolovius, 1804), pp. 8–9.

Religion and Reason

Kant's philosophy is usually seen historically as a retreat from religion and part of the *secularization* of modern culture under the sway of *Enlightenment reason*. Many features of Kant's philosophy support this picture. His critique was influential in discrediting the traditional metaphysical arguments for God's existence and the immortality of the soul. His moral philosophy replaces God's will with our own rational will as legislator of moral laws. Kantian aesthetics breaks decisively with the tradition that connects the experience of beauty directly with a divine order, tracing the experience of the beautiful instead to the harmony of our faculties of imagination and understanding. Kant treats experience of the sublime not as awe at the divine majesty but as our awareness of the way our own moral vocation transcends nature's power.¹

Kant also offered a moral argument for God's existence, but it is usually seen as an afterthought – a pitifully weak or even half-hearted attempt to repair what his critique had successfully destroyed. Kant's *Religion within the Boundaries of Mere Reason* is viewed by the religious as an attempt to reduce religion to morality, while secularists are unable to take seriously Kant's attempts to interpret Christian doctrines sympathetically. Both camps see Kant's treatment of Christianity in the *Religion* as a kind of Trojan horse within Christianity. Kant's "rational" religion is seen as secular morality appropriating religious concepts for its own purposes, while showing incomprehension of, and even contempt for, authentic religion.

As regards Kant's influence, it would be pointless to dispute this picture, since (like much that passes for "intellectual history") it is self-perpetuating and thereby self-verifying. We assert (tautologically) that Kant has had such-and-such influence because we understand him to have had that influence. The story

¹ Lewis White Beck speaks of Kant's "Copernican Revolution" as also a "Promethean Revolution," in which humanity usurps the place of God (Wood, 1984, pp. 28–30).

is then only about those who tell it; “Kant” becomes merely a placeholder for our prejudices about him. We omit the inconvenient fact that our narrative grossly falsifies Kant’s own self-understanding and the intended historical effect of his philosophy. It treats the common misrepresentations as true and interprets Kant’s real aims as if they were a false mask concealing his intent.

This book tries to present critically but sympathetically Kant’s project in the *Religion*. What Kant can teach us about religion and secularism in the modern world must emerge indirectly, as we think about what might be said for and against his project. I think it could teach us a lot, if we would let it, about the potentialities of religion in the modern world and the ways they have fallen short of being actualized.

1.1 WHAT IS “RELIGION”?

In Kant’s philosophy, the word “religion” (like many other words in his philosophical vocabulary) has a precise technical meaning: “Religion (subjectively considered) is the recognition of all duties as divine commands” (R 6: 84, 153, cf. KpV 5:129, KU 5:460, MS 6:487–488, SF 7:36, VpR 28:997–999). It is typical of Kant, however, that this statement needs clarification. The qualification “subjectively considered” means that “religion” refers to the attitudes of individual human beings. A *religious* person is one who thinks of moral duties as having been commanded by God. Religion “objectively considered” might refer to further divine commands belonging to *revealed* religion, whose relation to *rational* religion we will explore in this book. Kant insists that there is only one *religion* because the subjective attitude to which the term refers is the fundamentally same in all who regard their duties as divine commands. Kant distinguishes *religion* from a *faith* (*Glaube*) – more precisely, an ecclesiastical faith (*Kirchenglaube*). There are many ecclesiastical faiths, corresponding to different communities of faith, with differing beliefs in alleged divine revelations, on which these various communities have been founded (R 6:107–108).

In the literature on Kant, it is sometimes said that he means to “reduce religion to morality.”² This is false, and importantly false. For Kant there is no ethical duty for us to be religious – to think of our duties as commanded by God. But as we will see in Chapter 2, Kant holds that assent to God’s existence on practical grounds is something morality gives us reasons to do. In the preface to the *Religion*, Kant puts it this way: “On its own behalf morality in no way needs religion (whether objectively . . . or subjectively)” (R 6:3), but, on the other hand, “morality inevitably leads to religion” (R 6:6). Kant would deny that a morally good person has to be a religious person, though he argues

² Hare (1996), p. 45: “Kant and the Reduction of Religion to Morality.” Hare resists the common view that Kant does this; but as we will see later (Chapter 5, §3), Hare thinks Kant provides “unhelpful translations” of religious claims into moral ones. Cf. Palmquist (1992).

that morality provides reasons for being religious. The right way to describe the relation between morality and religion for Kant is to say that *religion goes beyond morality, adding something to it that enriches the moral life*.

One thing that clearly follows from Kant’s technical sense of the word is that “religion” for Kant is *theistic*, perhaps even *monotheistic*. This may look like a serious limitation if we think of “religion” as including polytheism, whether in the ancient world or in non-Western religious culture or in nontheistic traditions such as Buddhism. But it doesn’t mean that Kantian philosophy could not be related in interesting ways to these other religious traditions.³ It means only that they do not fall under the technical concept of “religion,” which it suits Kant’s purposes to use.

Kant normally assumes that a religious person believes there is a God. But he explicitly denies that this belief is strictly *required* for a person to be religious. A person can be religious as long as that person recognizes that *if* there is a God, then all ethical duties would be commanded by God (MS 6:487). Kant says that the “minimum of theology” necessary for religion is that the concept of God is the concept of a *possible* being (R 6:153–154 n). In his treatment of rational theology, Kant thinks we are capable of establishing that much, as well as offering ourselves a fairly detailed account of the attributes of such a possible being (VpR 28:998).

For Kant there is much more to being religious than thinking of God as commanding your duties. As we will see in later chapters, Kant thinks of religion as involving attitudes of awe and gratitude toward God. We will see that you must strive to be well-pleasing to God – seeking his forgiveness. Traditionally, God is the father of a religious community, regarded as a family.

³ Some argue that although Kantian philosophy and Buddhism share an emphasis on the awakening of our intellectual powers, Buddhist ethics has greater affinities to the Scottish Enlightenment’s emphasis on sentiment (see Cummiskey, 2017). This assessment, however, depends on a sadly common prejudice that stereotypes Kantian ethics as hostile to the emotional interconnectedness of human beings. This error assumes that reason excludes emotion, and it completely ignores such central themes in Kant’s ethics as the moral incentive: goodness of heart, love combined with respect, the realm of ends, and (in religion) the ethical community (see Chapters 4 and 7). Older scholars see affinities between Madhyamika Buddhism and transcendental idealism (see Murti, 2008), but these, too, strike me as spurious. A more pertinent comparison seems to me that between the Buddhist concept of “unanswerable questions” and Kant’s theoretical dialectic, especially the Antinomies (see Conze, 1963). But this book, the focus of which is on Kant’s thought about religion in his sense of the term, is not the place to explore these possible connections. As for polytheism, Kant does briefly discuss it (VpR 28: 1008, 1040). His account suggests that he would classify it as a form of what he calls “anthropomorphism,” one that assigns different departments of nature to different supernatural beings, represented according to human limitations. But Kant praises polytheistic cultures, both Asian and Western, for often seeing their many gods as expressions of a higher One Divinity from which they proceed.

1.2 RELIGION AS ESSENTIALLY SYMBOLIC

What does it mean to say that God is a *possible* being in the sense required for religion? For Kant there is an important distinction between merely *logical* possibility: the noncontradictoriness of a *concept*, and the *real* possibility of an *object*, which requires that our concept of the object have thinkable *content* – or, as Kant sometimes says, “sense and significance” (*Sinn und Bedeutung*). To think of God as commanding our duties, we must think of God as having an understanding and a will. For even the “minimum of theology,” therefore, we need a concept of God as a being with something analogous to human cognition and conation, hence analogous to the sensory content found in the concepts of objects we can cognize empirically. Since God is an idea of reason, not a possible object of empirical cognition, no sensory content is available from a theoretical standpoint. Kant thinks it can be supplied from a practical standpoint – through *symbols* (Ak 20:279–280, KpV 5:119–121, 134–141, KU 5:353).⁴

Religion provides us with a vocabulary – not merely of words but also of thoughts, feelings, images, narratives, and emotions – through which to experience our lives. For Kant this vocabulary consists of what he calls *symbols*. Religion is a way of thinking about our human existence in terms of symbols of the divine and also symbols relating our lives to the divine.⁵

That religion essentially involves symbolic thinking is a fact often neglected by both religious people and secular critics of religion. For both, the symbolic character of religion makes the content of religious belief harder to understand, and its neglect therefore makes life easier for them – but easy not in a good way. It makes religious beliefs easier for critics of religion to reject because when claims that are symbolic in meaning are treated as if they were not, then they often appear simply false, as when they come into direct conflict with scientific or historical truths, or seem to offer a view of the world that competes with science but is far less plausible when judged by rational standards. The same neglect, however, can make religious claims easier for religious people to accept because it facilitates their doing so dogmatically and uncritically, based on tradition, authority, and bigotry. It also seems to relieve them of the responsibility of interpreting the symbols of their faith. The failure to meet this responsibility inevitably impoverishes their religion. It often subverts or

⁴ We give the necessary experiential content to our thought of God by representing God as really possible through *symbols*. Chignell (2010b) appeals to symbolism as a way of addressing the problem that there might be a *Realrepugnanz* among the divine perfections that would render an *ens realissimum* really *impossible*. If Chignell is right, symbols are also a way of dismissing the thought that God might not be really possible. This cannot be done on theoretical grounds; it would have to be done on the practical grounds we will discuss in Chapter 2.

⁵ The Christian theologian who has made most profound use of the concept of religious symbolism is Paul Tillich. See Tillich (1951, 1952, 1957, 1963); cf. also Rowe (1968). This book is not the place for a detailed comparison of Kant and Tillich regarding the symbolic in religious thought. But we must not neglect to mention that Kant’s philosophy of religion is an important source of Tillich’s conceptions of religious faith and symbolism.

corrupts the human meaning of what they profess to believe. Literalism in religion risks turning truths into lies and good into evil.

For Kant the role of symbolism in religious thinking is a consequence of his most fundamental doctrines about thought and cognition.⁶ Kant's critique of reason famously denies that it is possible for us to have any theoretical or demonstrative knowledge of God's existence or any cognition of God's attributes. Cognition requires not only concepts but also intuitions – immediate cognitive contact with an object cognized through these concepts. Human intuition is sensible, but God is not a possible object of our senses, and we have no supersensible cognitive capacities that would permit us to achieve any cognition of God or his attributes. Our theoretical cognition can never represent God as anything but a logically possible concept of reason.

These limitations, however, do not preclude what might be called – and even what Kant himself might call – an “experience” of God. This is not in a sense of “experience” that refers to theoretical cognition but in one that refers to what we accept on practical grounds or as part of aesthetic experience. Such experiences are symbolic in character. *Symbolic* presentations of a concept correspond to *sensible* or *intuitive* ones but are importantly distinct from them. Theoretical cognition might arise from *schematic* presentations of a pure or a priori concept that enable the concept to be applied *directly* to empirical instances (KU 5:351, 352, cf. Anth 7:191). Ideas of reason (such as God or the ideal of human moral perfection), however, are pure concepts for which no direct or schematic presentation can be given. Symbols are essential to giving concrete or intuitive application to ideas of reason. All thinking about God as a really possible being, therefore, must be symbolic. “All our cognition of God is merely symbolic” (KU 5:353). (“Merely” here contrasts *symbolic* with *empirical* cognition.) When we represent our duties as divine commands, we are thinking *symbolically* about morality, simply because we are thinking about its relation to God.

Symbols are *indirect* sensible presentations, by way of *analogy*. Analogy for Kant is not a similarity between objects but a *sameness* (in the relevant respects) between the *way we think* about *different* objects (P 4:357, cf. KU 5:464, VpR 28:1023). Thoughts about our relation to God present the realities of our moral life to us indirectly, symbolically, in an aesthetically or emotionally significant way. But to think of an object by way of analogy calls our attention to our own thinking. It invites us to *interpret* our symbolic thoughts and the emotions they evoke. This means: to *reflect critically* on the thoughts we are having and *assume responsibility* for our interpretation of the symbols, rather than merely losing ourselves in our emotions. As beings who exercise rational agency, we must be both *feeling* beings and *thinking* beings.

⁶ The theme of symbolism in Kant is crucial for his philosophy of religion. This theme has been generally neglected in the literature. Some exceptions are Kang (1985), Bielefeldt (2001), Chignell (2006), and Pasternack and Rossi (2014). I discuss religious symbolism further in this book, both in this chapter and in Chapter 5 §§ 2–4.

Kant holds that it is one of the functions of “aesthetic ideas” to represent the supersensible. Aesthetic ideas are intuitive representations that suggest an inexhaustible association of thoughts that can never be adequately grasped in any determinate concept (KU 5:313–317). Kant illustrates aesthetic ideas by (pagan) examples of the sublimity and majesty of the divine: the eagle of Jupiter and the peacock of Juno (KU 5:314–316). Kant also associates the emotions occasioned by the feeling of the sublime with the deity venerated by Judaism, Christianity, and Islam; he argues that the strength of the religious emotions in these ancient monotheistic faiths depended precisely on the fact that this deity could not be an object of the senses in the form of a “graven image” (KU 5:274). The religious power of scripture, in Kant’s view, depends on understanding the words of scripture *symbolically* (R 6:110 n, 111, 136, 171, 176). Scripture often elicits emotional understanding of human situations by way of narratives and parables that move us but also cry out for interpretation if they are to apply *thoughtfully* to our lives.

Kant warns us that if we think about God in a way that is not symbolic, then we cannot think in a way that is genuinely religious. Either we lapse into “deism” (in Kant’s technical sense of the term) or fall into “anthropomorphism.” A *deistic* concept of God is a purely metaphysical concept of an *ens realissimum*. It is an idea of reason, a mere logical possibility, having no real content, no meaning for our lives and no religious significance for us. Religion must replace deism with “theism,” through symbolic representations of God. Far worse than deism from a religious standpoint is *anthropomorphism*. It thinks of God (superstitiously) as if he were a powerful yet personified natural force with power and will like a human tyrant (KpV 5:131, 136, KU 5:353, R 6:141, 168, VpR 28:1046). True religion is then replaced with slavish groveling and idolatry. Kant regards much of popular religion as susceptible to this abuse (VpR 28:1118).

As a natural and historical phenomenon, religion is as complex and ambiguous as human nature itself. The idea is still popular that there is such a thing as an irrationality that is above reason and that religion is it, or a part of it. This treacherous idea can be presented in entirely innocuous and even appealing forms, but Kant was well aware of its dangers. David Hume explored that dark side of religion in *Natural History of Religion*. For him, religion’s predilection for obscurity and unintelligibility is essential to it and closely allied to the ways it supports slavery, bigotry, and intolerance and even undermines morality (Hume, 1992). The history of every religion offers us powerful grounds to agree with Hume. Kant thinks, however, that Hume’s somber diagnosis captures only part of the truth. It can even be argued that Hume’s view of religion itself supports forms of religious irrationalism, by viewing them not as aberrations but instead treating the subversion of reason as the very essence of religion.⁷

⁷ For an account that raises this problem, see Kail, 2007. Kant’s friend Hamann, a religious antirationalist, seems to have read Hume’s view of religion in just this way. See Berlin (1977).

Kant's conception of religion is closer to that of Lessing, who views religion's vocation more hopefully as that of educating the human race (Lessing, 2005, pp. 217–240). Such a view enables us to see some varieties of religion as examples of true education and others as examples of miseducation. It can take account of the aspects of religion to which Hume calls attention, while at the same time offering a more hopeful picture of what religion has done for the human race and what it can still do for it. Lessing's idea that divine revelation can be seen as a source of human reason itself has affinity with the way Kant tries to understand the relation between revealed and rational religion that we will be exploring in Chapters 7 and 8.

For Kant it is only through symbolism that the pure concept of God can be presented in a way that is meaningful to human beings and therefore truly religious. It therefore gets things exactly backwards to speak of a religious concept or doctrine as “only” symbolic, if we intend by that to contrast a symbolic interpretation with some more “genuine” presentation of the divine, described as a “literal” one. On the contrary, it is the *merely literal* (i.e., the anthropomorphic) interpretation of a doctrine or scriptural text that is not genuinely religious because it is impoverished, inauthentic, corrupt, and superstitious. It substitutes the dead letter for the living spirit that the letter truly signifies. “The *ideal*,” Kant says, is then “mistaken for an *idol*” (Anth 7:192). Another Kantian term here is “delusion” (*Wahn*): the mistaking of a representation for what it represents (R 6:170–171). Like the Bible and the Koran, Kant recognizes *idolatry* (*avodah zera*, *shirk*) as both a permanent temptation to true religion and also a mortal danger to it. But it would be pointless to want to destroy the symbols, as if that could free us from idolatry in our religious attitudes. The harder path, but the necessary one, is to understand the symbols and assume responsibility for their morally and religiously enriching interpretation.⁸

When atheists reject religious doctrines, they usually do so on a literal interpretation. They then take them to be competing with natural science about the origin of the universe or of life, or with historical evidence about past events. So understood, religious doctrines are often fitting objects of

⁸ Symbolic thinking is not limited to religion, and although religion would like to think of its symbols as ways of experiencing good emotions, there is no question that evil emotions can also be powerfully symbolized through beliefs, sometimes even through literally true ones. Both Kristallnacht (which actually happened) and *Protocols of the Elders of Zion* (which was a forgery) symbolized hatred of Jews in Germany. The lie that Barack Obama was born in Kenya gave symbolic expression to the emotional resistance many Americans felt to an African American president. For many Americans, a border wall with Mexico is a symbolic wall separating them from a future they cannot accept. People often choose their heroes, whether Martin Luther King Jr. or General Robert E. Lee, as symbols of causes in which they still believe. Fixating on the literal object, or literal truth, of a symbolic belief is often a way of evading the question whether the emotions expressed symbolically are rational or irrational, good or bad.

intellectual contempt and even moral condemnation. What atheists reject in this way may be at least part of what many literalist (superstitious, anthropomorphic) believers actually believe. Literalism in religion is how many people perpetuate false archaic theories of the world and unreliable or merely legendary historical reports. It is how some subcultures use the authority of ancient scriptures to rationalize slavery, tyranny, racism, misogyny, and homophobia.

From a Kantian standpoint, these arguments (both pro and con) are not about religion but about anthropomorphic delusion or idolatry. To think authentically about true religion requires that we consider claims about God *symbolically*. No doubt it has been mainly philosophers and theologians who have raised the difficult questions about what religious symbols truly mean and how they are to be interpreted. Kant should be seen as this kind of philosopher. But it is a fundamental mistake to think that it is only philosophers or theologians for whom the meaning of words, concepts, and propositions used in a religious context is chiefly symbolic. In ancient times, people did not distinguish the literal meaning of religious thoughts from their symbolic significance. They seldom separated rudimentary scientific thoughts from symbolic religious thoughts. Ancient peoples also usually deferred to scriptures, traditions, and religious authorities without assuming responsibility for understanding the human (i.e., the symbolic) meaning of the thoughts they were accepting. This failure to think for themselves was an important part of the state of degrading intellectual minority from which human beings liberate themselves through the slow, never completed process Kant calls *enlightenment* (WA 8:35).

Kant thinks about the natural world in the ways we would regard as scientific. But he denies that science has the solution to all the problems of human existence. No doubt the motley collection of successful research programs we now call science can help us to understand ourselves better; many of them have already done so. But it is a far more doubtful wish that they, or perhaps some ideal combination of them, could ever suffice for us in dealing with all the absurdities and perplexities of the human condition. "Scientism" might be a pejorative term describing that ignorant, shallow, and humanly impoverished wish. One need not be hostile to science to think that human beings need other cultural resources besides science in dealing with the hard philosophical problems of consciousness, value, and freedom – as well as the existential perplexities of mortality, death, and grief; feelings of guilt and inadequacy; and hopes and fears – and in confronting the uncertain destiny of the things we care most about, including even the historical fate of the human institution of science itself. Kant regards religion as an important one of these resources. Both scientific and religious thinking might offer human beings the hope of deeper self-understanding. There is no reason why the two hopes would have to be competitors, still less, seen as mutually exclusive. It is a self-inflicted

spiritual wound of modern culture – both on the side of science and of religion – when they are seen that way.⁹

1.3 RELIGION IN KANT'S LIFE

Kant was a polymath. His interest in philosophy grew out of his study of natural science: physics, astronomy, and chemistry. He helped to invent the emerging sciences of physical geography and anthropology. His pragmatic anthropology was enriched by his wide reading of travel narratives and also imaginative fictional literature. Kant's theory of right exhibits detailed and subtle knowledge of Roman-based legal traditions and codes of his time. His writings on religion likewise display an acquaintance with scholastic and Lutheran theology. Kant understood his own philosophical outlook as religious.

Kant's relation to the religion around him was, however, troubled and ambivalent. His father was a poor saddler (or leatherworker). His parents, especially his mother, were devout Pietists. Pietism was a revivalist movement among German Lutherans that resembles other eighteenth-century revivalist religious movements, such as Quakerism and Methodism in England, and Hasidism among eastern-European Jews. If the eighteenth century was an age of reason and enlightenment, it was also an age of religious revival and emotionalism. The two were essentially connected. Fear of rationalist science and liberal society led many to retreat into traditional and emotional religious faith to protect themselves from the disorienting and alienating experience of modernity. Enlightenment rationalism itself was an encounter of modern thought with contemporary religious revival. In this respect, Pietism was typical of the age. It stressed literal adherence to biblical teachings and the experiential side of the religious life. A conversion or "born again" experience was regarded by most Pietists as essential to salvation. Pietism was also a proselytizing movement that emphasized the priesthood of all believers, the equality of all human beings as children of God, and the eventual hope for a church universal.

When Kant's Pietist pastor Franz Albert Schulz noticed exceptional intellectual gifts in the second son of the humble saddler, he arranged for Kant to be admitted to the newly founded Collegium Fredericianum. There Kant was not only exposed to Pietist zealotry but also prepared for a university education. In 1723, a year before Kant was born, Pietists in the court of

⁹ Kant no doubt belongs among the intended targets of Freud's impatient remark that "where questions of religion are concerned . . . philosophers stretch the meaning of words until they retain scarcely anything of their original sense" (Freud, 1964, p. 51). Freud, of course, has his own interpretation (grounded in his psychoanalytic theory) of the symbolism involved in many religious thoughts. When Freud sees religious thoughts as having symbolic psychological significance, he is engaging in the very same enterprise as the targets of the remark just quoted.

Friedrich Wilhelm I of Prussia had persuaded the king to exile Christian Wolff, professor at Halle, who was then the leading German philosopher of the day and the leading figure in the German Enlightenment (*Aufklärung*). Kant grew up in an atmosphere determined by this conflict. Many academic philosophers, such as Martin Knutzen, with whom Kant studied, attempted to reconcile Pietist religion with Wolffian philosophy. Kant began his university studies with Latin literature but soon came to focus instead on the natural sciences. He then turned to metaphysical issues concerning their foundation. Kant's concentration on ethics, politics, and religion came only later.

From quite early in his career, Kant was interested in the metaphysical concept of God and in defining the proper role that divine purposive design might play in the empirical investigation of nature. But it was only much later that religion, properly speaking – the moral and emotional significance of belief in God – came to be one of Kant's important philosophical interests. In the *Religion*, as we will see, Kant argues that an “ethical community” or church, founded on a revealed scripture, is necessary for the moral progress of humanity. But he also mounts strong philosophical and moral criticisms of existing religious beliefs and practices. In consequence of these criticisms, Kant refused on principle to attend services at the cathedral in Königsberg, even when his position as rector of the university included the expectation that he should.

During the reign of King Friedrich II of Prussia (Frederick the Great) (1740–1786), there flourished a relatively wide range of theological views both among the clergy of the established Lutheran church and among scholars and university academics. Some of them applied standards of modern scholarship to scripture and tried to harmonize faith with modern life. Prominent among the latter were Hermann Samuel Reimarus (1694–1768) and Gotthold Ephraim Lessing (1729–1781).¹⁰ Reimarus was a leader in the field of historical biblical scholarship called neology: the application to scripture of the same critical methods that historians had applied to other ancient historical documents. Reimarus concluded that the historical claims of scripture were largely lacking in evidential support. Lessing, who was responsible for the posthumous publication of Reimarus's *Wolfenbüttel Fragments* (1774–1778) held a more complex theological position, but one of

¹⁰ For two classic studies of these two thinkers, see Schweitzer (2005), chap. 2; and Allison (2018). Reimarus was also a representative of what was then commonly called deism – not in Kant's technical sense, explained above, but in the more frequently used sense of a belief in “natural religion” – that is, a theistic religion that dispenses with revelation or even rejects it. The title *Religion within the Boundaries of Mere Reason* might be taken to be “deism” in this sense. The title of my paper “Kant's Deism” (see Wood, 2002), if it is understood to claim that Kant is a deist in that sense, would then be asserting a falsehood, as I hope this book will show. But my article was meant to raise a question, not to assert a plain falsehood. (I've also written articles titled “Kant's Compatibilism” and “Kant's Historical Materialism.”)

its most prominent features was an insistence on religious toleration (conspicuously, of both Jews and Muslims, as well as different Christian sects).¹¹

The era of relatively free religious thought in Prussia came to an abrupt end in 1786 with the death of Frederick the Great and the accession of his nephew Friedrich Wilhelm II. The new monarch instituted a program of academic and clerical censorship under the administration of Johann Christoph Wöllner, whom Frederick the Great had described as a “scheming, swindling parson.” In 1788 Wöllner issued a new censorship edict, strengthening those that had been on the books (but seldom strictly enforced) during the reign of Frederick the Great. This met with resistance from the Superior Consistory of the Lutheran Church in Berlin, which had become relatively tolerant under Frederick’s reign. But Wöllner’s resolve was only stiffened by clerical opposition. Pastors suspected of unorthodox opinions were sometimes instructed to preach on certain texts (e.g., those involving miracle reports), and they lost their posts if they did not say the right things. Academic philosophers and theologians were subjected to a newly instituted Immediate Commission of Investigation. Those who had taught or written things disapproved of by the censorship were forced either to recant or to lose their professorships. Kant’s friend and colleague J. G. Hasse, a prominent neologist, retained his position only by recanting his earlier teachings, thereby destroying his scholarly reputation. Kant by 1790 was unlikely to be subject to such an extreme inquisition but might be faced with a similar dilemma if he were not cautious. For a proper understanding of the controversy in relation to Kant’s *Religion* and its publication, we must first introduce Kant’s project in that work.

1.4 THE EXPERIMENTAL TESTING OF A HYPOTHESIS

At the end of the preface to the first edition of the *Religion*, Kant described his project in the book as an “attempt” or “experiment” (*Versuch*) – or rather two of them. The first is “considering [biblical theology and pure rational religion] as a unity” (R 6:10). Thus the object of the *Religion*’s first experiment is to confirm the hypothesis that, as he puts it in the preface to the second edition, biblical religion and rational religion are related as concentric circles, and to refute the thought that the circles are merely overlapping, or even external to each other. Kant then proposes a second attempt (or experiment). He will start from some alleged revelation – specifically, the Christian – and to “hold fragments of this revelation, as a historical system, up to moral concepts, and

¹¹ Lessing’s alleged adherence to the views of Spinoza was to be the subject of a sensational controversy in 1785–1786 between F. H. Jacobi (1743–1819) and Lessing’s close friend, Moses Mendelssohn (1729–1786). See Chapter 8 § 2. Kant entered the controversy on Mendelssohn’s behalf with his essay *What Does It Mean to Orient Oneself in Thinking?* (1786).

see whether it does not lead back to the same pure rational system of religion” (R 6:12).¹²

Before we look at Kant’s experiments themselves, we should first consider carefully the hypothesis they are supposed to test. Many people nowadays, whether religious or not, may find Kant’s hypothesis difficult to entertain it at all. The revealed religion Kant is proposing to use for his experiments is the Christian religion. People from other religious traditions could presumably consider an analogous hypothesis for their revealed faith. Kant’s hypothesis rests on a single basic background assumption, which sometimes steps into the foreground. It is that both rational and revealed religion have it as their fundamental aim to make better human beings of us. As religious people, subject to the commands of a perfect God, our basic end is our own moral improvement. It is on this that our eternal welfare must ultimately depend. If religion is essentially the recognition of our duties as divine commands, then whatever might be valuable about any religion must be valuable because it somehow contributes to our being the kind of people who would try to obey these commands or adopt other attitudes that belong to moral virtue, making us better human beings or (as Kant puts it) human beings “well-pleasing to God” (R 6:62). Religion should encourage feelings and attitudes of joy, gratitude, and hope because these attitudes are often part of our becoming good people or better people. Kantian religion, like Kantian ethics, does not identify moral goodness with happiness. Religion is not there only to offer us consolation or make us feel good irrespective of our efforts to become better human beings.

According to Kant’s hypothesis, then, pure rational religion and revealed religion constitute two concentric circles, both having the religious vocation of making us into better human beings. The hypothesis then contains three main elements or claims:

First, there is such a thing as pure rational religion. In other words, it is possible, on merely rational grounds, to regard our moral duties as commanded by God. This requires us to believe both that we have moral duties resting on rational grounds independent of divine revelation, but also that we can think of these same duties as commanded by God.

Second, the inner circle of pure rational religion lies *entirely within* that of empirical divine revelation. No rational duties directly conflict with duties recognized by revelation, and all rational duties can also be regarded as belonging to those recognized by revealed religion. If you are a Christian, you cannot think that parts of rational morality might not also be binding on you *as a Christian*. That requires that you interpret Christianity so that Christian revelation is *compatible* with rational morality. There can be no “teleological suspension of the ethical” – at least if that dubious Kierkegaardian phrase refers to valid

¹² Pasternack (2017b) reaches the startling and paradoxical conclusion that there are not two experiments but no “experiment” at all. He fussily translates *Versuch* as “attempt.” The second *Versuch*, he thinks, refers only to the second edition of the *Religion*. But it is evident that Kant is testing a hypothesis, which makes the word “experiment” at least as appropriate a translation as “attempt.” And the second edition clearly identifies *two* such attempts (or experiments).

divine commands directly to violate duties valid in rational ethics. If, for instance, revelation seems to say that God commands a father to murder his innocent son just as he might slaughter a sheep, then that portion of scripture must somehow be interpreted in a way compatible with the rational judgment that so monstrous an act is of course morally forbidden (R 6:87, 187). But use of rational argument to support an *interpretation* of revelation is in no way a *rejection* of the revelation.¹³

Third, those who accept a rational religion in this sense can see Christian revelation as in some sense adding something more to what reason can strictly justify. There might be revealed doctrines that reason cannot know or even accept on rational grounds, and there might also be duties recognized by revealed religion that fall outside the duties of rational morality. But the revealed doctrines must be interpreted so that they do not contradict what we know or rationally believe, and the revealed duties must not include actions reason rightly judges to be immoral.

Some doctrines of revealed religion, therefore, belong also to rational religion, while others do not; but rational religion does not deny those doctrines of revealed religion that it does not include. We might describe these same relations between the outer circle of revealed religion and the inner circle of rational religion using a terminology Kant offers us in the introduction to the *Critique of the Power of Judgment*.¹⁴ In the third Critique, Kant thinks of any set of concepts or principles as constituting a *field* (*Feld*). The part of this field within which cognition is possible for a given faculty is called that faculty's *territory* (*Boden, territorium*); and the part within which the faculty is *legislative* is called its *domain* (*Gebiet, ditio*) (KU 5:174). In these terms, the doctrines and duties belonging to pure rational religion constitute both its *territory* and its *domain*. Some doctrines and duties belonging to revealed religion also fall within both the *territory* and the *domain* of rational religion. But some further doctrines and duties belong only to revealed religion; they must fall within the *territory* of rational religion, since their content must be rationally cognizable; but they fall outside its *domain* because they are legislated not by rational religion but only by revelation. Conversely, the doctrines and duties belonging to revealed religion must at least encompass those of rational religion, since they constitute the outer or wider concentric circle of religious faith. They must be part of the *territory* of revealed religion and also belong to its *domain* to the extent that revealed religion specifically includes them within its revealed legislation. But there will also be doctrines and duties falling outside the *domain* of rational religion that do belong to the *domain* of revealed religion.

¹³ Taken literally, therefore, the title of Savage (1988) – “Kant’s Rejection of Divine Revelation” – can express only a falsehood. If Kant had *rejected* divine revelation, then his entire project in the *Religion* would have collapsed. What Savage actually claims (correctly) is only that Kant denies empirical divine revelation counts as *theoretical cognition* (Savage, 1988, p. 61). But to deny revelation this philosophical status is not necessarily to reject it. Kant not only considers revealed religion a permissible form of religion but also regards some of its doctrines as symbolizing what can be accepted by rational religion. As we will see in Chapter 7, he even holds that revealed religion is a historically necessary vehicle for rational religion, so that rational religion is dependent historically upon revealed religion for its very existence and even its content.

¹⁴ This is done by Moors, 2017, though not necessarily in just the way I am about to do.

Of course both rational and revealed religion must reject certain doctrines, or the apparent teachings of certain scriptures, when these are given certain (anthropomorphic or *merely literal*) interpretations: for instance, the claim that a good God could literally command a man to slaughter his own innocent son like a sheep. About the story of Abraham and Isaac, Kant says that an acceptable interpretation may “often appear to us as forced, in view of the text (of the revelation), and may be forced in fact; yet if the text can at all bear it, it must be preferred to the literal interpretation that either contains nothing for morality, or even works contrary to its incentives” (R 6:110). However, Kant does not presume to usurp the function of the biblical theologian in interpreting revealed doctrine. He argues on rational grounds for some interpretations of Christian doctrine over others but never sets himself up as a rival to ecclesiastical authorities for the purpose of declaring what is orthodox Christianity.

Kant advocates toleration of all religious convictions; thus there is no place in Kant’s religious thought for concepts such as “heresy” or “anathema” (cf. R 6:108–109).¹⁵ Some religious sects, of course, do engage in plainly wrongful *practices* – human sacrifice, child molestation, deprivation of women’s rights over their own lives and bodies, burning people at the stake for following their consciences. These terrible practices are violations of right and should be coercively prohibited by any just state. But they are not “anathema.” Consider this analogy: Some uses of property are criminal, but the theory of property right does not need a category of “criminal property right”; likewise, revealed religion does not need a category of “anathema” to treat of wrongful *practices*. That they are *wrongful practices* is all that matters. But doctrines, however mistaken, and even when they are held in ways contrary to morality, are never *wrongful*. They *never* deserve malediction, execration, and coercive suppression.

1.5 THE FIRST EXPERIMENT

Kant’s first attempt or experiment (*Versuch*) is to see whether a “religion,” a recognition of all duties as divine commands, is possible within the boundaries

¹⁵ Pasternack claims that Kant regards some things in Christian revelation (Original Sin as biologically inherited, the Incarnation, Vicarious Atonement, Abraham’s sacrifice of Isaac) as “anathema” (Pasternack, 2017a, p. 532). I have to wonder whether Pasternack understands the meaning of this ugly word; if he does, then this must be hyperbole, perhaps even a (misleading and tasteless) joke. The word “anathema” implies coercive prohibition of a doctrine, even an execration or curse upon it – something Kant would never pronounce on any doctrine, since doctrines are parts of free human communication. For Kant, even to *reject* any Christian doctrine would be inconsistent with treating revealed religion as an outer circle *surrounding* the inner circle of rational religion. It would be as self-contradictory as saying that Tompkins County is within New York State but that other New York State counties (say, Schuyler, Steuben, Allegany, and Cattaraugus Counties) are not within the same state as Tompkins County.

of mere reason. “Mere” (*bloß*) (sometimes translated, or mistranslated, as “bare” or “naked”) here carries the sense of *unaided* (sc. by divine revelation), whether through scripture, tradition, or other empirical sources. The hypothesis to be confirmed is that when the teachings of revealed faith are properly interpreted, there is no conflict between them and the religion of pure reason (MS 6:488). The experiment is harder to carry out for Kant than it might be for those who think God can be cognized theoretically, either through pure reason or in some other way. Kant does, however, offer an argument, based on moral considerations, that we have rational practical grounds for acceptance of God’s existence. The argument is briefly summarized right at the beginning of the preface to the first edition of the *Religion* (R 6:3–6). We will be examining it in Chapter 2.

Hume’s Philo thinks *how great God is* must be an object of merely verbal dispute, like arguing over which effusive compliments to pay the beauty of Cleopatra (Hume, 1980, p. 80). For Kant it is nothing of the kind. For Kant, if there is a God, he must be a supremely perfect being, not merely a *very great* or *very powerful* being. Unless God is supremely perfect, there is no ground for symbolizing his will as perfectly good, and therefore no reason for thinking that his commands, if there are any, would coincide with our moral duties (VpR 28:1118). This also means that the idea of God is not an empirical concept but a concept of reason or an idea, to which no experience could ever be adequate. In the *Critique of Pure Reason*, Kant considers how we acquire the concept of God – of a supremely perfect being or *ens realissimum* (KrV A571–583/B599–661). He argues that this concept is an *idea* in a technical Kantian sense of the term: it is a concept to which no possible experience could ever correspond (KrV A310–338/B366–396). We have this idea not because we do, or even could, have experience of its object but because our rational faculty infers it from concepts of finite things according to rational procedures. In his *Lectures on the Philosophical Doctrine of Religion*, Kant gives a fairly extensive account of the contents of a concept of God, that correspond to the concept found in scholastic and early modern philosophers (VpR 28:1012–1116, cf. KrV A567–583/B595–661, KU 5:438–453). But of course Kant thinks there can be no *theoretical* grounds, rational or empirical, for claiming cognition of the existence of such a being.

It is basic to Kant’s entire ethical theory that moral duties themselves are not grounded on belief in God (R 6:3–4). The foundation of ethics – the material end on which its categorical imperatives are based – is the dignity of humanity or rational nature as an end in itself (G 4:427–429). This is objective, holding independently of our contingent and fallible wills and equally independently of the necessarily good will of God. Kant could not be more explicit about this: “The essence of things does not alter through their external relations, and it is in accordance with that which alone constitutes the absolute worth of the human being, without thinking about such relations, that he must be judged by whomever it may be, even by the highest being” (G 4:439). Nevertheless, once

we conceive of God as a supremely perfect being (an *ens realissimum*) and then symbolize God as a perfect will, consistency requires that we think of God as willing the performance of all genuine duties.

Clearly more is required to recognize duties as God's *commands* than merely acknowledging duties and believing, assenting, or hoping that God exists. We must somehow think of God (again symbolically) as an *authority* commanding us to fulfill our duties. Kant's moral argument for God's existence contains some reasons for us to have this further thought. For when we think of God as proportioning happiness to worthiness or otherwise doing *justice* in the world, we think of God in the way we think of an ideally just ruler of a human community doing justice among its subjects. The formula of the "realm of ends" treats humanity as an ideal community in which all rational beings are *members*, but a supremely perfect rational being is thought of as its *supreme head*: "But a rational being belongs as a *member* to the realm of ends if in this realm it gives universal law but is also itself subject to these laws. It belongs to it *as supreme head*, if as giving law it is subject to no will of another" (G 4:433).

These would be reasons for having the highest respect for God, even the highest love. We have a duty to display attitudes of respect and love toward all rational beings (MS 6:448–452, 462–465, ED 8:337–338; see Chapter 4 §2); in relation to God we should be moved in addition to do our duties in part by an attitude of veneration toward the supreme head of the ideal moral community or realm of ends (R 6:7–8, cf. G 4:433–434). It is also crucial, however, that we *not* obey God's commands only in order to gain heavenly rewards or avoid divine punishments. That would turn obedience to duty into mere hired or compulsory service (*Frondienst*) (R 6:158, 165, 171, 180, ED 8:339, MS 6:484, VpR 28:1084). Kant distinguishes between "fear of God" – a disposition to obey God's commands from respect for law and from love of God as a holy will (R 6:182, cf. 6:145, KU 5:448–9 n) – and "being afraid of God." The latter is a slavish and even an immoral attitude (KU 5:260, cf. VE Collins 27:1425–1426, VE Vigilantius 27: 556–557, 725–726).

1.6 THE SECOND EXPERIMENT

The referent of the title of Kant's book *Religion within the Boundaries of Mere Reason* is the specific way of recognizing our duties as divine commands, which falls within the domain of reason unaided by empirical revelation. In parts one and two of the *Religion*, and also in Chapters 3 through 6 of this book, religion is being considered from the standpoint of the individual moral agent, even the individual sinner who is seeking salvation. Its basic concepts are moral ones, though they are also symbolized religiously through the individual's relationship to God. In parts three and four of the *Religion*, however, Kant turns to that side of religion that involves a religious community and, more specifically, the way religion is actually present in the real world, in churches or ecclesiastical faiths that base their service of God on claims to empirical

revelation, as preserved in an ecclesiastical tradition and in scriptures or holy books. In Chapter 7 we will see that for Kant, once this social aspect of religion is taken into account, the religion of reason cannot exist independently of ecclesiastical faith – nor, conversely, can true religion in ecclesiastical form exist entirely separate from the religion of pure reason. The two forms of religion must then be seen as standing in a historically dynamic relationship, in which rational religion exists through the symbols of an ecclesiastical faith, and ecclesiastical faith, too, is to be interpreted, and also reformed, through reason. This ongoing historical dynamic is anticipated in Kant’s “second experiment.”

In part two of the *Religion*, we are to consider the grounds for hope that we may become “well-pleasing to God” through a “moral revolution” or “change of heart,” struggling against the propensity to evil we find in ourselves. In part three, Kant argues that this struggle cannot be carried on successfully by each of us alone, but requires that we unite in a free ethical community or “people of God.” Both aspects of pure rational religion are clearly modeled on the fragments of Christian revelation Kant is considering in his second experiment. The striving to be well-pleasing to God is related to the Christian doctrines of salvation and grace, and the ethical community is modeled on the Christian church. The second experiment involves “starting with some alleged revelation” and “abstracting from the pure religion of reason (insofar as it constitutes a system on its own), to hold fragments of this revelation, as a historical system, up to moral concepts and see whether it does not lead back to the same rational system of religion” (R 6:12).

The starting point for the second experiment is some “alleged” divine revelation.¹⁶ We might wonder whether, consistent with Kant’s philosophy, divine revelation is even possible. For Kant’s theory of knowledge denies that we could ever have empirical cognition of a divine being – the object of a mere idea of reason. We could never be empirically acquainted with anything reason could confirm as coming from God. The limitation here, of course, is ours and not God’s. Kant does not think we could ever be in a position to deny that a divine (an omnipotent) being could produce effects on us of whatever sort it chose. But our limited faculties of knowledge would be incapable of cognizing those effects as divine in origin. As he puts it: “If God should really speak to a human being, the latter could still never know that it was God speaking” (SF 7:63). Based on Kantian epistemology, to assert dogmatically that a certain experience you have was an experience caused by God is to assert something directly contrary to the laws of your own human understanding. Thus to claim with confidence that you have received an empirical divine revelation is necessarily *superstition* in Kant’s technical sense of the term: using your

¹⁶ One quite thorough discussion of the question where in the *Religion* the “second experiment” is to be found is Firestone and Jacobs (2008), pp. 114–119. As is by now evident, however, I reject their claim that the second experiment begins only in part four of the *Religion*.

understanding in a way that is contrary to its own essential laws (KU 5:294). It is equally superstition to assert that you have experienced some special event (a miracle) that was caused by God (R 6:53).

It follows that no empirical or historical credential could ever establish that a holy scripture is divinely inspired (R 6:187, SF 7:64). The most reason could confirm is that the *content* of the alleged divine speech is such that, as far as we can judge, it might have come from a supremely wise and morally perfect being. Kant does think, however, that we can reasonably judge that something claimed to be divinely revealed is *not* genuine, if what is supposedly revealed is *contrary* to reason or the moral law (SF 7:63, cf. R 6:87, cf. Fichte, 2010, §§7–14, pp. 64–131). Kant even argues that we may regard the moral law as a kind of “inner revelation,” which is the touchstone of any alleged “outer” revelation through experience (VpR 28:1117).

The closest we can come to knowing that something taught or reported in an alleged holy scripture was actually revealed by God is that *as regards its content*, we can *rationally judge* that this content *might have been revealed* to us by God. This turns out to be a thesis of more positive significance than we might have supposed because in part three of the *Religion*, Kant also argues that it is a tenet of pure rational religion that our moral improvement can be achieved *only* through membership in a moral community or voluntary commonwealth, seen as a “people of God” under moral laws (R 6:95–100). He argues further that in human history the only way that moral communities of this kind have ever been founded is that they have come about as *churches*, based on some alleged revelation preserved in scriptures (R 6:100–107). So it turns out that pure rational faith is driven by its own principles to take seriously various claims to empirical divine revelation that churches have made. It is not an option, therefore, to reject revealed religion wholesale or simply to throw away parts of scripture if you don’t happen to like them. For Kant accepts the necessity of revealed religion as part of his project.

Kant has a set of technical terms to describe the various positions one might take regarding the relation of rational to revealed religion. A *naturalist* is someone who accepts a pure rational religion and denies that there is any revealed religion. A *supernaturalist*, by contrast, is someone who thinks revealed religion is universally necessary for any genuine religion. In between these two positions is one Kant calls that of the *pure rationalist*: it allows for the possibility of divine revelation but claims it is not necessarily required for religion (R 6:154–155).¹⁷ Kant clearly wants to identify his own position with

¹⁷ In “Kant’s Deism” (Wood, 2002, p. 98), it seemed to me not to make sense to admit the possibility of divine revelation but then to suggest we pay no attention to it. Lawrence Pasternack has correctly observed that pure rationalism *does* make sense because revealed doctrines can be thought of as religiously permissible and even religiously helpful for some people as long as the revealed doctrines are interpreted in a way that makes them serviceable to a pure moral religion (Pasternack, 2014, pp. 13–14, 219–223). Of course he’s right: that is precisely Kant’s position.

pure rationalism. Pure rationalism does not deny the possibility of divine revelation, even if it does deny that we can ever know for certain that some alleged revelation is genuine. By allowing for the possibility of revelation, pure rationalism departs from the strict naturalist position. And because for Kant even pure rational religion is committed, for its own purposes, to consider the alleged revelations in the scriptures of historical churches, it makes sense to ask how far the alleged revelation of some particular church might, if rationally interpreted, lead back to the religion of reason. In other words, it makes sense for Kant to conduct his second experiment using some specific alleged divine revelation. Living where and when he does, Kant naturally chooses Christianity, even recognizably Lutheran Christianity, as the revealed religion in terms of which the experiment is to be conducted. It would be unreasonable and anachronistic to expect him to have proceeded in any other way.

In its second experiment, then, Kant's *Religion* considers "fragments" of an alleged revelation – namely, the Christian. Specifically, it takes up these: Original Sin, the Son of God as savior, divine grace, and the church. The hypothesis Kant is assessing in his second "experiment" is that revealed religion is a wider concentric circle surrounding the inner circle of rational religion. To confirm his hypothesis, Kant must offer an interpretation of Christian doctrine that is at least consistent with pure rational religion. The Christian doctrines found in the outer circle cannot be *rejected* by rational religion, though they are also not embraced or included within it. Kant's position on these doctrines is that rational religion should regard them as religiously optional. They are religiously acceptable, but they do not belong to a "saving faith" – a faith making the individual worthy of eternal happiness, which, Kant says, can be only one faith, despite the diversity of ecclesiastical faiths (R 6:115). They might be helpful to religion, however, if the symbols are given an appropriate interpretation.

We will see in Chapters 7 and 8 that what belongs to the inner circle of rational religion is not to be viewed as a set of beliefs and practices that are now fully determinate and static. Instead, it refers to the telos of an ongoing historical process of religious reinterpretation and religious reform. For it is through this process that revealed religion – the outer circle – shows itself to be the vehicle of a pure religious faith whose very historical possibility consists in the existence of this vehicle and in its historical self-reform through rational enlightenment. It can therefore be only provisionally that Kant can characterize the referent of the title of his book. Perhaps he can refer to it in the form of a provisional criterion for distinguishing those interpretations of symbolism in revealed religion that exhibit it as a vehicle of pure religious faith. But this criterion itself can be only provisional. As we will see in Chapters 7 and 8, the line separating the two concentric circles – of pure religion and revealed religion – is meant to be shifting and even to blur over time as the dynamic process of religious enlightenment and development continues. If Kant sees the two concentric circles as still distinct, this is because the religious reforms for

which he hopes are only now beginning. If we still see the circles as sharply distinct, then we should recognize this as a sad reminder that Kant's hopes for enlightened religion and human moral progress have been disappointed. That failure belongs not to Kant but to modernity and to institutional religious faith (specifically, to Christianity).

1.7 THE PUBLICATION OF KANT'S *RELIGION WITHIN THE BOUNDARIES OF MERE REASON*

The hypothesis Kant's second experiment is attempting to confirm would, if wholly true, show that *everything* in revealed (Christian) faith is compatible with rational faith. It would disprove the hypothesis if there are parts of revealed Christianity that are demonstrably and permanently incompatible with rational religion. But the second experiment attempts to confirm the hypothesis only for selected fragments of Christian revelation. The aim of Kant's experiment is also to interpret revealed Christianity in such a way that the hypothesis can be confirmed and show that reason and Christianity are allies, not enemies. If this cannot be done, then all Christians who are also rational moral agents would find a hopeless conflict at the very center of their spiritual lives. Kant's project in the *Religion* is to show them how this can be avoided.

It was widely known that as soon as he completed the *Critique of the Power of Judgment* (1790), Kant turned his attention to religion. Thus, when in the spring of 1792 there appeared, under anonymous authorship, a short book titled *Attempt at a Critique of All Revelation*, it was hailed by its earliest reviewers as the awaited contribution by Kant. In fact, however, as Kant soon hastened to acknowledge, the book was the very first publication by the hitherto unknown J. G. Fichte (1762–1814). Kant's plan at this point was to publish his work on religion in the form of a series of essays in J. E. Biester's journal *Berlin Monthly*.¹⁸ Biester had, however, already moved this journal to Jena, outside Prussian jurisdiction, obviously in order to avoid the newly instituted censorship, and because there was no prohibition on the importation of books to Prussia from the Grand Duchy of Saxe-Weimar where Jena was located.

Since Kant was a Prussian academic, teaching in Königsberg, his essay on religion needed approval from the appointed censor. This was G. F. Hillmer, a secondary school teacher and notorious religious bigot, assisted by J. T. Hermes, an equally extreme local pastor. (It is indicative of the difficult political position Wöllner's censorship efforts faced in the context of the resistant Lutheran ecclesiastical hierarchy that in order to carry them out, he

¹⁸ The following account of the history of Kant's publication of the *Religion* is based on Dilthey (1890), pp. 418–50. A briefer account in English is given by George di Giovanni in *Ca Religion and Rational Theology*, pp. 41–54. See also Di Censo (2012), chap. 1.

needed to enlist such incompetent subordinates.) Kant's first essay for Biester's journal, on the radical propensity to evil in human nature, was approved by Hillmer, probably because (as we will see in Chapter 3) it proposed a version of the Christian doctrine of Original Sin. It was published in the *Berlin Monthly* in April 1792 (about the same time, in fact, as Fichte's book). Kant's second essay, however, dealing with more directly biblical issues relating to divine grace was not so fortunate. In the spring of 1792, it was refused Hillmer's *imprimatur*. Biester attempted to appeal this decision to higher authorities in Berlin, but to no avail.

Kant anticipated even greater problems from the censors with the third and fourth essays he was planning, which were even more directly critical of existing churches, including the Lutheran. He therefore devised an alternative plan. His writings on religion were no longer to appear as a series of essays in Biester's journal, but were to be published together, as four parts of a single book; and this book was to be published in such a way as to avoid having to be approved by the likes of Hillmer and Hermes. For an academic work, one legally acceptable route to its publication was approval by a relevant university faculty. In the case of Kant's *Religion*, it was unclear whether this would be a theological faculty or a philosophical faculty. Kant therefore asked the theological faculty at the University of Königsberg for its opinion whether his work fell within the province of the theological or the philosophical faculty. Essentially, he was asking the Königsberg theologians to waive jurisdiction over the work on behalf of theological faculties generally. In the event, Kant got the waiver he sought. He then submitted the entire work to the philosophical faculty at the University of Jena, located in the Duchy of Saxe-Weimar and therefore beyond the authority of Prussian censorship. The Jena philosophical faculty approved it, and *Religion within the Boundaries of Mere Reason* was published by Nicolovius of Königsberg in spring 1793. A second edition, with a new preface and many additions indicated by a dagger (†) was subsequently published in January 1794.

The Berlin censors were outraged by what they saw as Kant's devious way of circumventing their authority. As early as June 1791, Kant had already been apprised by his student Kiesewetter (a royal tutor in Berlin) that some sort of punitive action against him was coming (Ak 11:264–266). Two of Kant's writings indicate what his response would be. The second part of *On the Common Saying*, published in Biester's *Berlin Monthly* in September 1793, defends the position that a subject is obligated to obey even the unjust commands of a legitimate ruler. *The End of All Things*, appearing in June 1794, was (in Kant's own description) "partly plaintive and partly funny to read" (Ak 11:496–497). It argues that Christianity's highest claim is to be a religion of love and then argues that the spirit of Christian love is a *liberal* spirit, incompatible with coercion in religious matters (ED 8:336–339). This is clearly Kant's anticipatory protest against the impending royal actions against him.

In October 1794 Wöllner wrote to Kant in the name of King Friedrich Wilhelm II. “Our most high person,” his letter said, “has long observed with great displeasure how you have misused your philosophy to distort and disparage many of the cardinal and basic teachings of the Holy Scriptures and Christianity . . . particularly in your book *Religion within the Boundaries of Mere Reason*, as well as in your shorter treatises” (SF 7:6). The letter in effect forbade Kant henceforth either to lecture or to write on religious topics. In reply, Kant presented an unyielding and rigorously argued defense of his conduct. He denied that the philosophical project undertaken in his writings offered any *appraisal* whatever of Christianity; it therefore could not possibly have been guilty of *disparaging* the Christian faith. Kant further pointed to the many ways in which his book both directly and indirectly evidenced respect for Christianity and its holy scriptures (SF 7:8–9). Kant then concluded by offering the king this solemn promise: “*as Your Majesty’s most loyal subject*, I will hereafter refrain altogether from discoursing publicly, in lectures or writings, on religion, whether natural or revealed” (SF 7:10).

The italicized portion of the promise (Kant’s emphasis) turned out to be especially significant. For as he made explicit in a footnote when he later published this letter, Kant took his promise to be one directed specifically to King Friedrich Wilhelm II. Therefore, when that king died in November 1797, Kant took himself to be released from its obligation. Less than a year later, he published *Conflict of the Faculties*, which included not only a further discussion of religion and biblical theology but also a defense of both the legality and the propriety not only of his earlier writings on religion but also of the route through which the *Religion* was published.

1.8 WÖLLNER’S CHARGES AND KANT’S DEFENSE

We may have a hard time understanding the context of Kant’s actions surrounding the publication of the *Religion*. We all believe (or suppose ourselves to believe) in freedom of speech, freedom of publication, academic freedom, religious freedom. Americans especially may even reject the very concept of an established church. We are therefore likely to see the censorship laws and practices within which Kant was forced to operate as an outrageous violation of basic human rights that we take for granted. On the other hand, if we imagine the position of an academic subject of the eighteenth-century Prussian monarchy, we may then also suppose that there is no point in a subject claiming, on some subtle legal pretext, the right to evade this (blatantly unjust) power. Kant’s arguments in favor of his actions may therefore strike us as “legalistic” in a bad sense of the word. In the language I quoted from Lawrence Pasternack in the Preface, Kant’s reasoning in support of his actions are apt to smack of “guile” and “prevarication” (Pasternack, 2014, p. 12).

In order to understand the issue between Kant and the Prussian censors, however, we need to appreciate that although he believes in the legitimacy of the

authority of Friedrich Wilhelm II, and thus even accepts an obligation to obey even his unjust commands, he understands all of this in the context of his own well worked-out theory of political right. This theory grants to a ruler the right to limit freedom of expression in many ways of which we would probably not approve; but it does not treat this right as unconditional or unbounded. It places particular stress on the right of free expression of academic writers, and most especially of members of the *philosophical* faculty of a university. We need therefore to look briefly at Kant's conception of a university, and its basis in political right, a topic which is also of some interest for its own sake.¹⁹

According to Kant's theory of right, any executive authority administering a legitimate condition of right is charged with the coercive protection of the rights of all citizens (MS 6:311–338). In order to have the capacity to do this, a ruler must do certain things to look after the welfare of the populace that it can reasonably demand. These include care for people's physical, civil, and eternal welfare. The state must supervise the training of three professions – physicians, lawyers, and clergy – and consequently certify their competence. This training and certification is the job of universities, which the state must therefore establish and support. The three faculties charged with these tasks are the three so-called higher faculties of medicine, law, and theology. They are “higher” in respect of their nearness to state authority. Each must establish legally binding standards for members of the corresponding profession (SF 7:17–23).

In order to do so competently, however, the state must depend on a fourth “lower” faculty, the faculty of philosophy, whose task is to investigate the objective truths on which the state authority bases the statutory standards set by the three “higher” faculties (SF 7:23–27). The faculty of philosophy, however, can have no similar statutory constraints, since its function is not to follow the legal authority of the state but simply to determine objective truth (which of course is not determined by state authority). The members of this faculty must of course meet standards of scientific competence within their fields, but it would be counterpurposive to try to lay these down in advance by legal statute (SF 7:27–29). Moreover, Kant sees the philosophical faculty as the center of that “learned public” whose free give-and-take is the condition of the enlightenment of human beings generally (WA 8:35–42).

On Kant's theory, these are the considerations that should determine whether a given piece of writing that touches on both philosophy and some “higher” faculty properly belongs to the jurisdiction of the one faculty or the other. When Kant asked the theology faculty of the University of Königsberg whether it was the proper authority to which to submit his book for censorship, he was in effect asking whether the contents of *Religion within*

¹⁹ For a fuller account of what I am about to say, see Wood (2012), which, however, is in German, since it was part of an inquiry into the historical background of present-day reflections on the place of the university in German culture and society.

the Boundaries of Mere Reason were to be judged by the authoritative standards governing the training and certification of clergy or instead only by the standards of objective truth pertaining to a philosophical faculty. The Königsberg theological faculty quite reasonably ruled that philosophical and not theological standards applied to this particular book. Consequently, Kant was officially permitted to submit his book to a *philosophical* faculty, which he then did: at the University of Jena, which approved it. Again, Jena was not in Prussia but in the Duchy of Saxe-Weimar. However, books approved there could be legally sold in Prussia.

By the standards recognized in Prussian law and the Prussian university system, therefore, Kant was not engaging in “guile or prevarication” in asking for and accepting the rulings and permissions through which the *Religion* was approved for publication. There is no “legalism” in Kant’s defense, except in the sense that the publication of the *Religion* followed the laws valid for that time and place. In short, in his defense Kant was appealing simply to the rule of law. If the king, or his minions, thought that he was violating the proper jurisdiction of their censorship, then they were *wrong on the law*.

Further, when Wöllner’s letter of reproof declares that Kant’s *Religion* had *appraised*, and also *disparaged*, Christianity, that was *wrong on the facts*. In his reply, later published in *Conflict of the Faculties*, Kant correctly denies he had done either. Kant proposes to defend a *religion within the boundaries of mere reason* – that is, a way of recognizing our duties as divine commands that is not dependent on any alleged revelation, including the Christian (R 6:3–6). Such a project surely does not in itself appraise Christianity, still less disparage Christianity. Kant’s image of the concentric circles implies that revealed (Christian) religion goes beyond the religion of mere reason, but also that it does not contradict it. Further, the hypothesis Kant means to confirm is that certain prominent fragments of Christian revelation *do* lead back to pure rational religion. This argument proposes to confirm, as by experiment, the hypothesis that at least regarding those fragments, the two circles are concentric. If successful, therefore, Kant’s experiment could not disparage Christianity, but would in fact do the direct opposite. In his reply, Kant insists that the only object up for appraisal in his book is *natural* religion, the religion of mere reason, considered apart from any revealed faith. The aim is to judge how far religion within the boundaries of mere reason can defend the tenets of *its* faith (SF 7:8–9). Kant also insists that the *Religion* evidences “great respect” for Christianity through its claim that the Bible is the best available means for popular instruction in morality and religion (SF 7:9–10).

The *Religion* might be seen as containing, at least by implication, a “disparagement” of Christianity, but only *if* we make two assumptions: *first*, that the interpretations of Christian doctrines Kant offers involve the direct denial of important Christian truths; and, *second*, that agreement with the religion of reason, as Kant presents it, is a standard that any purported revelation must meet. The *first assumption* would be highly dogmatic. Even

state- or church-authorized biblical theologians, not to mention free rational inquirers (such as members of the philosophical faculty of a university that is not bound by church authority), might justifiably disagree with it. Kant is not proposing to tell either the established Lutheran church or the theologians of the “higher” theological faculty how they are to interpret Christian doctrines. He is merely offering his own interpretation of these doctrines: an interpretation that enables them to be reconciled with a religion of reason.

Nor is the *second assumption* – that agreement with rational religion, as Kant presents it, is a standard revealed religion ought to meet – anything Kant would presume to impose on theologians or the established church. Kant might accept this standard himself, and he might try to present it in such a way that it would appeal both to philosophers and to Christians. But it would be up to each individual reader of Kant's *Religion* to decide whether Kant has correctly presented the truth about morality or a religion of reason, and also what role rational religion should play in that individual's own thinking about religion or about life.²⁰

Of course if I am a Christian reader of the *Religion* who finds himself making both of these assumptions, then that would create a problem *for me*. It would suggest that I cannot bring the beliefs of *my own Christian faith* into harmony with the deliverances of *my own reason*. But it would still be a mere distraction for me to ask, as many writers still do, whether Kant “accepts” or “rejects” this or that part of “Christianity.” If I disagree with Kant's philosophical arguments about rational religion, or with his interpretations of Christian doctrine, then I should present my case to myself on rational grounds and decide whether Kant's philosophical claims and interpretive suggestions are right or wrong. Kant is presenting for his readers' consideration what *his* reason tells *him*; he is not claiming to dictate what *yours* must tell *you*.

The tacit assumption, however – sadly made by many in all ages – is that we are not answerable to our own reason in matters of religious faith. We are entitled – perhaps even required – to impose on ourselves, perhaps even on others, certain irrational dogmas, or even to defer to some dogmatic authority in preference to – even contrary to – our own reason, concerning what we ourselves should think and accept. That assumption was presupposed by the Prussian authorities who condemned Kant's teachings. It is directly opposed to what Kant calls “enlightenment” – the struggle to *think for oneself* (WA 8:35).

Too many present-day scholars in effect make the same assumption when they focus on the question whether Kant “accepts” or “rejects” this or that

²⁰ In *Conflict of the Faculties*, Kant clearly explains the difference between legitimate and illegitimate conflicts between the theological faculty and the philosophical faculty (SF 7:32–67). Pasternack (2014, pp. 71–81) completely ignores this explanation when he claims falsely that Kant is violating his own rules for separating philosophical from theological consideration of Christian doctrines. He apparently thinks Kant violates these rules simply by using reason to suggest interpretations of Christian doctrine that would reconcile rational with revealed religion.

Christian doctrine. That focus must also always rest on some dogmatic assumption about what Christian doctrine is – that is, about how it should be interpreted; and it must dismiss all rational argument as irrelevant to how Christian doctrine ought to be interpreted. These scholars look upon Kant’s philosophical arguments as dogmatic pronouncements and therefore treat his interpretation of Christianity as if it pronounced “anathema” on any alternative interpretation. This approach thereby dishonestly evades the only real question raised by Kant’s thought about the Christian religion: How should Christianity be understood if we are to reconcile its doctrines with our own reason?

Moral Faith in God

2.1 PHILOSOPHICAL ARGUMENTS FOR BELIEF IN GOD

Arguments in support of belief in a deity are easy to dismiss and even easy to ridicule. Unbelievers naturally find it insulting for a philosopher to pretend by some clever argument to effect the radical alteration not only in belief but in life-orientation that would be required for them to come to believe in God. For believers such arguments are insulting in a different way, since they seem to assume that their faith is so weak and shallow that it might need support from some clever philosopher's twist of metaphysical reasoning.

For many believers, moreover, thoughts about God are apt to be considered in the context of religious practice and judged by the standards believers apply to acts of worship.¹ Philosophical arguments of any kind, however, especially in the field of metaphysics, are usually attempts to work out some kind of philosophical perplexity. As applied to God's existence, they are often attempts to resolve a quandary about the scope of necessary truth or the structure of causality or about purposiveness in the natural order or in biological explanation – or, in the case of Kant's arguments, about a problem of practical reason regarding rational

¹ Some theistic proofs are meant to function at the same time as acts of worship. These belong to the “faith in search of understanding” tradition, which would include not only arguments offered in various works by St. Augustine (author of the phrase “fides quaerens intellectum”) but also St. Anselm's famous (so-called ontological) argument, which occurs in the context of a prayerful reflection on his own attitude of faith, in the first four chapters of *Proslogion*. Jacques Maritain treated Aquinas's Five Ways as expressing consciousness of our cosmic dependency. Arguments from design can also be presented in a spirit of worship directed at the wisdom and goodness of the creator, such as the argument given by Hume's Cleanthes. Do such arguments lose philosophical credibility on this account? I tend to think so. But not all theistic proofs need be offered in this pious spirit. Most are not, and despite its engagement with our moral predicament, Kant's moral argument is not.

assent to the possibility of its ends. Arguments of that nature, if they are honest, will be offered in a critical mood of tentativeness and even skepticism. This can strike a believer as irreverent or impious. But if a theistic argument is not presented with due critical modesty, its philosophical credibility may be drowned in the baptismal river of pious emotion. There seems an inevitable tension between religious piety and philosophical honesty. Even if the theistic argument is offered confidently, it can appear to introduce the Almighty into our scheme of things only as a means of solving some philosophical riddle – thus reducing the infinite Deity to the pitifully diminutive size that befits so petty a human purpose. According to Gordon Michalson, “Kant’s moral argument induces what we might characterize as ‘theistic shrinkage,’ generated by the requirements of practical reason. The God of the moral argument is chiefly an instrument in the realization of a rational goal and little more” (Michalson, 1999, p. 21). Kant himself attacks “religious fetishism” for reducing God to “a mere means to produce an effect in the world for which [the human being’s] own powers alone . . . do not suffice” (R 6:177). Michalson in effect accuses Kant’s moral argument of doing precisely that.

How ought we to think of the intent of Kant’s moral argument for belief in the existence of God? Few such philosophical arguments aim to convert unbelievers; most try only to offer those who already believe some intellectual articulation of why they believe. This is true of Kant’s moral argument. He says that it is “not intended to prove to the doubter that God exists; rather, it is meant to prove that if his moral thinking is to be consistent, he **must include** the acceptance of this proposition among the maxims of his practical reason” (KU 5:450–451 n). Not to accept the argument does not make you wicked or even necessarily irreligious. At most it makes you inconsistent in your practical attitudes toward the final end of morality when this end is conceived systematically from a philosophical standpoint.

The background of all Kant’s thinking about God is a metaphysical conception of the divine nature as *ens realissimum*. A *theistic* concept of the Deity includes symbolic representations of the attributes of understanding and will, insofar as these can be made consistent with supreme ontological perfection.² Kant famously rejects the traditional theoretical arguments for the existence of such a being, but his own theoretical philosophy makes extensive use of the concept of an *ens realissimum* for regulative purposes, both in metaphysics and in the philosophy of nature, while insisting that this use does not amount to theoretical *cognition* of God (KrV A631–704/B659–732, A825–827/B853–855, cf. KU 5:436–442). The philosophers and theologians in Kant’s audience would also have taken that

² This is presented in Kant’s lectures on rational theology and also integrated into his theoretical philosophy in bk. 2, chap. 3, of the Transcendental Dialectic of the first Critique (KrV A567–614/B595–642, VpR 28:1012–1117). My contribution to the scholarship on this topic is found mainly in Wood (1978, 2009). A critical review of the literature that accuses Kant of “atheism” or of reducing God to a mere “fiction” or “thought-entity” can be found in Palmquist, 2000, pp. 434–439.

conception of God as a rational philosophical representation of the being believed in by traditional monotheistic (Jewish, Christian, Islamic) religion, hence the being supposedly revealing itself in the holy scriptures of these faiths. Kant even devotes considerable argumentation to establishing the proposition that only belief in a God conceived according to the traditional metaphysics could answer to the needs of practical reason (KrV A810–811/B838–839, KpV 5:127–129, KU 5:442–447, VpR 28:1011–1012, 1071–1075).³ Kant always takes the philosophical issues he is addressing to be about the existence of a traditional scholastic-rationalist Deity. Kant’s moral argument does not aim to revise that concept of God. Whether or not it is successful, the God to whose existence it aims at justifying assent is no more “shrunkened” than the supremely perfect God of Aquinas, Scotus, Descartes, or Leibniz.

2.2 PRACTICAL GROUNDS OF ASSENT

We need to begin by understanding clearly what Kant means by a “moral” argument for God’s existence. Some arguments for belief in God that have been called moral are really theoretical rather than practical, at least as Kant would understand the distinction. For example, a divine command theory of morality, such as those advanced by Robert Adams and Philip Quinn, might be used to show that the best account of *what moral truth consists in* would commit us to the existence of a morally commanding God (Adams, 1987, pp. 97–122, 144–164; Quinn, 1978). Richard Swinburne has argued that belief in God is needed to explain our *awareness* of moral truths (Swinburne, 2004, p. 218). But if the truths that allegedly constitute evidence for God’s existence are truths *about morality*, that does not make these “moral” or “practical” arguments in Kant’s sense of the term.⁴ There is the same kind of confusion in Michalson’s suggestion that Kant’s moral argument really depends on the assumption that “the universe makes sense” and that “the believer does not assent to belief in God in order to be moral, but only to satisfy reason’s demand for symmetry [between happiness and worthiness]” (Michalson, 1986, pp. 501–515). If Kant’s argument were as Michalson here describes it, then it would first require some *theoretical* ground for asserting the “symmetry,” and then Kant would need to provide a *theoretical* reason for believing that the existence of God is necessary to explain the symmetry. Thus Michalson seems not to understand the kind of argument Kant is offering.

³ Michalson’s charge might better be understood as rejecting that argumentation. But that would not justify Michalson’s conclusion that Kant’s is a “shrunkened” God. Claiming that a philosopher’s argument is a failure does not entitle you to attribute to *him* some conclusion it pleases *you* to draw from the alleged failure. If Michalson thinks Kant’s moral argument is *successful* but proves the existence only of a “shrunkened” God, then Michalson should take that unwelcome conclusion upon himself and not foist it on Kant, who obviously rejects it.

⁴ It is misleading, therefore, to say (as some have) that Kant is arguing for theistic belief based on “moral evidence.” See Palmquist, 2000, p. 77, n. 14.

A second way of misunderstanding Kant's conception of a "practical" ground for assent would be to see it as offering us some *inducement* to *do* things to make ourselves believe. Suppose I were offered a large sum of money in case I somehow made myself believe, for example, that seven is not a prime number. Pascal's famous wager, for example, seems to be an argument of that kind.⁵ Pascal claims, based on probabilistic reasoning, that there is a practical advantage to being a Christian believer. But he clearly sees that this is an argument not for believing but for *doing* certain things – taking holy water, having Masses said, acting as if he believed – which he then hopes will *cause* him to acquire the religious beliefs he would find advantageous – but for which beliefs he seems to know very well that he has *no good reason*. Once we see the distinction between (1) having an instrumental or prudential reason to *do something to bring about a belief in ourselves* and (2) having a reason for *the belief itself*, we can see that, properly speaking, *there are no instrumental or prudential reasons for beliefs themselves*. "Prudential reason for belief" is an oxymoronic or even a nonsense expression.

Kant's argument does not offer us any *benefit* or *advantage*, not even a moral or religious advantage, if we believe in God. It does not say you should believe something because belief is more likely to get you to into heaven or even likely to make you a morally better person. If someone offers you instrumental reasons to take actions that might cause you to deceive yourself in ways you find advantageous, that offer is morally corrupt; your accepting it would be equally deplorable whether the supposed advantage is financial, religious, or even moral.⁶ Any benefit or advantage to you of believing is *never* by itself a reason for belief or sincere assent to anything. At most it might be a sign that such reasons are lurking somewhere in the vicinity – as when a belief is advantageous because it is *true*, and there are theoretical arguments or evidence for this.

Kant's moral argument, by contrast, proposes to offer a *genuine reason* for honest, sincere inward assent to God's existence – an assent based on distinctively *practical* considerations. Once we understand the kind of argument Kant is offering, we should also have a response to Peter Byrne's contention that Kant's claim that "knowledge that God exists is not possible . . . also rules out the possibility of faith, where this entails believing or thinking that God exists" (Byrne, 1979, pp. 333–335). For we will see that practical considerations *can* rationally justify an assent to the proposition that God exists, which, however, is not knowledge of any kind and never could be converted into knowledge.

The basic idea underlying Kantian practical faith is a fundamental distinction between two kinds of desire or volition: the distinction between *wish* and *choice*.

⁵ Pascal (1995), §233. This is the way that people heavily invested in the tobacco industry or heavily invested (financially or politically) in the fossil-fuel industry manage to hold beliefs contrary to the obvious facts. Upton Sinclair was right: it is hard to get someone to understand something when their income depends on their not understanding it (Sinclair, 1934, p. 109).

⁶ Thus I clearly misrepresented Kant's moral argument when I said some years ago that it shows that belief in God would be "advantageous [for a moral agent] to have" (Wood, 1992, p. 404). By itself the fact that a belief is advantageous is *never* a reason to believe or assent to anything.

It goes back at least as far as Aristotle (Aristotle, *Nicomachean Ethics* III/2 1111b7–30). If we find some object attractive but do not intend to take any action to bring it about, then we may *wish* for the object but we do not *will* it. In merely wishing we do not set the object of our wish as an end of action or choose to do anything we regard as a means to it (G 4:394, KU 5:177 n, Anth 7:251). We merely wish for an object when we either do not regard ourselves as having any means tending to bring about the desired object or when we do not want to put forth effort to get what we desire. But to *will* an object as our end is to lay down for ourselves a rational norm: *to do something to bring it about*. If the application of that rational norm is to make sense, then we must take ourselves to have some *means* to do something that at the very least would *tend* to bring it about (G 4:427). Consequently, we must conceive of our end as at least possible of attainment and moreover at least partly through the actions we take toward it.

These features of purposive action ground a rational relation between the practical act of setting some object as an end and the theoretical assent to certain propositions involving its attainability. We need not regard our action toward the end as *certain* to achieve it; but we violate a norm of reason – involving ourselves in a rational incoherence between our actions and our theoretical assents – if we set some object as an end while declining to assent to the claim that the object is at least *possible* of achievement and that our actions are at least *possible* contributions to bringing it about. Conversely, if we do set some object as an end, we are then rationally committed to assent to its possibility and to there being actions of ours that could in some way contribute to making it actual.

It might seem that this relation between purposive action and assent could work in only one direction. We might suppose, namely, that in order rationally to set some object as an end, we must always *first* establish through theoretical evidence or argument that it is possible and that some act of ours would tend to make it actual. Only *then*, we might think, could it be rational to set it as an end and take actions as means. But this supposition holds only in the case where theoretical grounds alone either make it *certain* that the object is impossible or else make it *certain* that we are powerless to contribute to its actuality. In cases where the possibility of an object and our possible contribution to its actuality remain theoretically uncertain, it would still be rational to set the object as an end, as long as we also accept a *rational commitment to assent for practical purposes* to the object's possibility and to our being able to make some contribution to making it actual. In that case, it would be rational for us to accept these commitments at least partly on *practical* grounds arising out of our purposive volition. That crucial insight is the core of Kant's moral argument for God's existence, stated in its most concentrated form.⁷

⁷ My first attempt to explain and defend this kind of argument is found in Wood (1970, 2009). I still think the account I gave was largely right, but both it and my subsequent presentations of it stand in need of development and even of correction on certain points (for instance, see n. 6 above). In this chapter I am attempting to provide these clarifications and corrections.

As Kant puts it, “Once an end is proposed, the conditions for attaining it are hypothetically necessary” (A823/B851). The necessity, however, is only “subjective,” and assent to the necessary conditions is only “subjectively sufficient” (A822–82/B850–851). That is, I acquire a rational ground for theoretical assent to the possibility of an object as an effect of my action *only because I have chosen to set the end*. Setting the end is up to me because it is always an act of freedom (MS 6:381): it is the decision to impose on myself the norm that *ceteris paribus*, my actions should serve the end. That free choice makes this ground of theoretical assent “subjective,” even though it is also a *rational ground* for me to be *convinced*, not merely a *psychological cause* of my being (nonrationally) *persuaded* (A820/B848).

It might be supposed that this would have to be mere “wishful thinking” and not genuine rational assent. That supposition would clearly be wrong. For this kind of assent can play an indispensable role in rational action. A clear and accessible case of this is what Kant calls “pragmatic belief” (A824/B852). Suppose a physician examines a patient suffering from some respiratory symptoms. It is her professional responsibility to do something to help the patient, if she possibly can. The physician does not know with certainty that the patient’s condition is hopeless or that she can do nothing to improve it. In this case there are good *practical* reasons – the patient’s need, the physician’s professional responsibility – for her to set the end of improving the patient’s condition. In order rationally to choose actions toward this end, however, she must assent to the claim that helping the patient is at least possible. Further, to choose a definite course of action, she must assent to some hypothesis about the correct diagnosis and treatment of the patient’s disease. It is a misunderstanding to view this assent as merely useful or advantageous. No; it is absolutely *required by practical reason* for the setting of an end and the choice of means. Kant assumes the physician does what she does in part on the basis of empirical evidence, but the uncertainty of the situation is such that the theoretical evidence alone does not justify any particular diagnosis or treatment. It must be in part for practical reasons, therefore, that the physician assents not only to the possibility that the patient’s condition can be improved but also to whatever diagnosis provides the basis for her treatment.

In Kant’s example, the physician assents for these *pragmatic* reasons to the diagnosis that the patient has *consumption* (*phthisis*) and thus chooses those actions that would count as the most effective treatment for this ailment. Her assent is then fully rational, despite the fact that there are no adequate *theoretical* grounds for it. The grounds for her assent are also *subjective* grounds because they are based on her free choice to set the end and take means to it. For suppose a second physician has the very same knowledge and expertise but no responsibility to treat this patient. He might propose a different diagnosis, or – because the theoretical evidence is mixed or weak – he might simply suspend judgment about what is ailing the patient and even about whether the patient’s condition can be effectively treated. For this second, merely *observing* physician, the *practical* ground for the *treating* physician’s diagnosis would not exist. This

makes the treating physician's grounds for her assent both *subjective* and *practical* but nevertheless *rational*; they are based on her free but rational choice to set the end of helping the patient. Her assent to the diagnosis, and to the possibility of helping the patient, involves really assenting to them and doing so on clearly rational grounds, even in a case where knowledge is not possible. Kant's moral argument for assent to the existence of God deploys these very same considerations, and, contrary to the contentions of Byrne quoted above, it likewise aims to justify genuine rational assent even in the absence of any possibility of knowledge. But there are also certain crucial differences. These differences constitute Kant's reason for distinguishing what he calls *moral belief* from what he calls *pragmatic belief*.

Belief (capitalized). We should ask whether the assent that is produced by practical arguments – that is, what Kant calls *Glaube* (belief, faith) – is really what we ordinarily think of as *belief*. I think that in certain respects it is not. But it is certainly a fully *rational assent to theoretical judgments*. In this book I am going to follow Andrew Chignell's practice and use the term "Belief" (capitalized) to refer to this special kind of rational assent on practical grounds (see Chignell, 2007, pp. 324–325). In §§ 8–9 of this chapter, we will ask what sort of assent Belief is, after we have reviewed the different forms taken by the moral argument in the three Critiques.

Kant maintains, for reasons we will consider presently, that all moral agents are rationally required to include among their ends a comprehensive moral end which Kant calls the *highest good*. In §4 we will look further at what this end is. For now it is best to think of it only as the comprehensive end of morality, systematically conceived, and also an end that Kantian ethics maintains that every rational agent has a moral duty to set. Kant also holds that we can know which actions we ought to take toward this end – namely, those that fulfill our moral duties. But it cannot be *known* (through evidence or theoretical proof) whether this end is possible of attainment through those actions; and we can be quite certain that the end cannot be achieved *solely* through actions that are in our power. This element in the moral argument is of a purely theoretical character, but it is also entirely negative, consisting merely in the *absence* of theoretical grounds for holding that the highest good is either possible or impossible. Kant also maintains, however, based again on theoretical grounds – concerning the nature of the highest good as we must conceive it – that if the highest good is possible, this could be only because there exists a supremely perfect – omniscient, omnipotent, and morally perfect – author of nature, a God. In order rationally to set the highest good as an end, therefore, we must assent to the existence of God, since God's existence is a condition of the possibility of the highest good.

The possibility in question is *real* possibility, not merely *logical* possibility. There are many things that are logically possible (e.g., traveling faster than the speed of light, leaping tall buildings at a single bound) that it would not be

rational for anyone to set as an end. This means that where the conditions of their possibility transcend our cognitive powers – for instance, involving the power and will of God – then this real possibility must be supplied by practical rather than theoretical reason. As already mentioned in Chapter 1, the real possibility must also be presented *symbolically*. Below we will think of such a presentation as *conceiving of* the possibility of our end, or as making it *practically meaningful or comprehensible*.

From this we can see how the moral argument has the same basic *practical* structure as that of pragmatic Belief. In both cases, theoretical assent becomes rational for someone when that person has set an end while being theoretically uncertain of its possibility. In the case of moral Belief, however, unlike that of pragmatic Belief, the obligation to set the highest good as an end is not contingent or discretionary, but morally obligatory and universal: according to Kantian ethics, every moral agent is required to set this end. Further, all moral agents are in the same position regarding the theoretical claim that connects the end with its possibility of attainment. Whether there is a divine author of nature is theoretically uncertain independently of any contingent empirical considerations, such as those that led the physician to select a particular diagnosis. Those two points make the practical ground for assent to God's existence into a universal rational ground of assent, which, however, is still both practical and subjective because setting the highest good as an end is something we might choose or not choose to do. A person who did not set the highest good as an end might be subject to blame if this displayed a moral vice (and not merely a theoretical error or deficiency). Either way, that person would not have the practical ground for assent that agents have when they do set the highest good as an end.

2.3 WHAT IS AT STAKE IN KANT'S PRACTICAL ARGUMENTS

It is an indication of the importance of the moral argument for Kant that he presented a version of it near the end of each of his three Critiques. He also presented it repeatedly in other writings, such as the essay *What Does It Mean to Orient Oneself in Thinking?* (1786) as well as in his lectures on philosophical theology and in the first preface to the *Religion*. There are significant differences between different versions of the moral argument – especially between the first Critique version and the later ones. §§ 5–7 of this chapter will present the argument as it appears in each of the three Critiques.⁸ I have already begun by

⁸ The moral argument was historically crucial to Reinhold's getting the *Critique of Pure Reason* accepted as an important development in philosophy (see Reinhold, 2005; and Beiser, 1987, pp. 232–236). Adickes (1920) claimed that in the *Opus Postumum* Kant repudiated the moral argument based on his claim that the ground of our belief in God is the thought of God as moral legislator. For a recent defense of Adickes's thesis, see Kahn (2018). But there is no *refutation* of the moral argument in the *Opus Postumum*, not even a wholly explicit *repudiation* of it. See Wood (1970, 2009), pp. 11–13, to which, even after fifty years, I have nothing to add.

presenting some basic ideas that are found in all the different versions and will continue this presentation in this section and the next. I will end the chapter by asking precisely what kinds of propositional attitudes Kant's arguments justify regarding the claim that God exists.

A basic element in all forms of Kant's moral argument is that we are morally required to set the highest good (*summum bonum*) as an end. It is best, however, *not* to begin by focusing on the highest good, for at least two reasons. *First*, the idea of the highest good in Kant's philosophy is itself a complex, even an ambiguous or protean idea, which takes different forms and plays different roles in different contexts. For example, Kant thinks of ancient Greek ethics, in all its schools of thought, as founded on the idea of a highest good, regarded as an ideal of how to live (see Wood, 2015). It may refer at least partly to this idea even when used in Kant's moral arguments. But the role of the highest good is not even the same in all three forms of the argument presented in the three Critiques. As I've said above, for now it is best to think of it as simply the concept of the most encompassing end of moral action. *Second*, however, focusing on the highest good right at the start would inevitably distract us – as it has most discussions of Kant's moral arguments (including my own) – from the larger significance of this kind of argument. This is a significance that transcends the Kantian doctrines involved in his moral argument and even transcends the aim of justifying Belief in God. Kant's is a *kind* of argument that should matter even to people who are not the least bit religious and who are utterly unconcerned with the existence of God. That is the point I want us now to appreciate.

In 1908 the great American philosopher Josiah Royce published a book titled *The Philosophy of Loyalty* (Royce, 2008). In it he argued that human lives have meaning only because people can devote themselves to something larger than self. His general name for this larger thing was "loyalty." I must admit I distrust that word because in my experience – as an American, living in the late twentieth and early twenty-first centuries – loyalty is usually demanded of us by institutions that do not deserve it: corporations, professions, universities, political parties, the American political state and its military; or else by leaders who defraud, mislead, exploit and degrade those who are foolish or corrupt enough to be loyal to them.⁹ "Tribal" loyalties are humanity's worst enemy. Such loyalties impoverish our lives rather than giving them meaning.

Royce was fully aware of this, and therefore spoke of "evil" or "predatory" loyalties. Loyalty to vulnerable individuals, which is something in which I can believe, usually takes the form of protecting them from predatory loyalties. But loyalty to such an individual is always also loyalty to something larger and to something social. Loyalty to a spouse or partner is loyalty to a love relationship; loyalty to a friend is loyalty to friendship. Even if those things that most

⁹ Lord Acton (1834–1902) famously said, "Power corrupts; absolute power corrupts absolutely." He also said, "Great men are almost always bad men" (Dalberg-Acton, 1907).

intrusively demand our loyalty do not deserve it, there are nevertheless other things, even very large things, which do deserve our loyalty. Royce spoke in this connection of “universal” loyalty and even of *loyalty to loyalty itself*. He meant that whatever it is to which you choose to be loyal, you can be truly loyal to it only if you are also loyal to the idea of loyalty. Even if (like me) you dislike the word “loyalty,” surely you must admit that Royce was right in substance: the meaning of our lives depends on loyalty, in his sense of the term, to something larger than self.

Royce calls the object of any loyalty a *cause* of some kind. He is especially interested in loyalty to what he calls a “lost” cause. By this term he does *not* mean failed causes, still less deservedly failed causes (such as the South’s attempt to preserve slavery), or even hopeless causes, those that we can be quite certain will never succeed. By a “lost” cause Royce means any cause that cannot be fulfilled within the lifetime of the loyal community or any of its members. He even thinks of a “lost” cause in this sense as the only possible source of true and absolute norms for finite human communities. Obviously it is not irrational to be loyal to causes that are “lost” in this sense of the term. In fact, as Royce thought, our lives can have meaning only if there are some lost causes to which we can be loyal.

Clearly the *highest good* in Kant’s conception of it – as an end that is a duty for each of us, and a shared end for the ethical or religious community as Kant conceives of it (R 6:97) – can be seen as a “lost cause” in precisely Royce’s sense. To state Kantian doctrine partly in Kant’s and partly in Royce’s terminology, we have a moral duty to be loyal to this lost cause. My thesis in this section, however, is that the basic idea of Kantian practical faith need not be framed in terms of loyalty to the highest good. In fact, to grasp that basic idea, it may even help if for a bit we stop thinking about the Kantian *summum bonum* and let ourselves consider loyalty to any cause in Royce’s sense – “lost” but not hopeless. We also do not need to conceive of our commitment to this cause as a moral duty. We might instead think of it simply as a lost cause that gives meaning to our lives, even if our devotion is not thought of as a moral duty.

Secularists often devote themselves to lost causes in Royce’s sense. They work for the eventual triumph of the ideals of a free and just society, or of liberal democracy, or of a socialist society beyond capitalism, or peace on earth, or the liberation of some oppressed group, or merely for the long-term survival of the human species on earth, which is now mortally threatened by human-made climate disruption. I have devoted much of my own professional career to the lost cause of advancing the work of science and scholarship. Happy indeed is anyone who has devoted him- or herself to a worthy lost cause; pitiful indeed are those whose lives have served no lost cause but have instead been squandered shortsightedly on merely selfish ends, such as the contemptible quest for power, wealth, honor, or pleasure.

If you agree with me about this, however, then you must admit that anyone who has led a meaningful life is committed to something recognizably parallel

to Kant's moral argument. For all loyalty (in Royce's sense) involves something like a practical faith – if not in God, then in a social movement or in nature or in history or in the human species. For none of these lost causes is guaranteed success, but we who are loyal to them must set such a lost cause as an end. Anyone who is committed to ends that transcend their own lives and efforts therefore has compelling reasons to be interested in Kant's moral argument for God's existence. For this argument offers a model for the kinds of practical grounds for assent that are necessarily connected to any loyalty to a lost cause. Secularists might also do well to consider the possibility that Kant's moral theism, or theistic Belief more generally, could be ways of *symbolizing religiously* some of those very same lost causes to which these secularists are also loyal under secular conceptions of them.

2.4 THE HIGHEST GOOD AS A MORAL END

The idea of the highest good appears many times in Kant's philosophy. Even the term "highest good" (*summum bonum*) is ambiguous, referring sometimes to one idea and sometimes to another. Many different yet often overlapping or intertwining reasons are given why the idea of the highest good is of interest to philosophy, and there are also different accounts given of the reason why every moral agent is obligated to make the highest good an end.¹⁰ In this discussion I cannot hope to untangle or give order to these accounts, and I also do not need to do this. In §§ 5–7 I will try to say what the term "highest good" seems to mean in each of the three Critiques and how it motivates the moral argument for the existence of God.

One thing that seems to remain constant in Kant's conception of the highest good is his commitment to the heterogeneity of the human good – specifically, the distinction between two kinds of good and the relationship of priority or conditionality between them. The supreme good (*bonum supremum*) is the *moral* good, virtue, goodness of will, conformity to duty or the moral law. Distinct in kind from this is the *natural* good, well-being or happiness, which is also good (and intrinsically good, not merely good as a means to something else) but good only on the condition that it is related to morality. This relation is not always conceived by Kant in quite the same way. Happiness is not truly good except on the condition of worthiness for it. It is even something evil when it consists in, or is achieved by means of, something contrary to morality. It is happiness proportioned to worthiness that constitutes the conditioned component of the entire highest good. Consequently, the highest good, in the sense of the *complete* good (*bonum consummatum*), consists in morality combined with happiness proportionate to worthiness.

¹⁰ For discussions of the highest good in different works, including all three Critiques and the *Opus Postumum*, see Guyer, 2005, pp. 277–313; and also Mariña, 2009, pp. 329–355.

One thing that sometimes remains unclear is whether this combination of morality and happiness is meant to refer to the highest good of a single person – the moral agent to whom the argument is addressed – or instead to the highest good of an entire world. The former concept would apply to *you* if *your* happiness were proportionate to *your* worthiness, while the latter would be found only in an entire world in which *everyone* enjoys happiness in proportion to individual worthiness. As I mentioned earlier, in Kant’s treatments of the history of ethics, the concept of a highest good or *summum bonum* is more often understood as an individual human ideal: the *kind of person* you should strive to be (VE Collins 27:247, VE Vigilantius 27:483, VE Mrongovius 29:599). There is no doubt that in the moral argument, Kant often presents the highest good to us in this way, as an idea to which we should strive to make our individual lives correspond. Each of us is to ask, *What ought I to do?* – and then also, *For what may I hope, if I do as I ought?* (KrV A808–809/B836–837). Kant invokes the ancient conception of the highest good as a virtuous and at the same time happy life in arguing for the heterogeneity of the good and the conditional relation of the goodness of the moral to the goodness of the natural good. Kant criticizes the Epicureans and the Stoics for identifying virtue (the moral good, the good of our person) with happiness (the natural good, the good of our state or condition) (KpV 5:111–113). Instead, he argues that the highest good is “**the human being** (each rational being in the world) **under moral laws**” (KU 5:448).

There is also no doubt, however, that at times Kant presents the highest good as an object of collective pursuit by the human species as a moral community (R 6:97, Anth 7:333). It is the highest good of an entire world: a “moral world” (KrV A808/B836), “the highest good of a possible world” (KpV 5:111), the “highest best world” (*höchsten Weltbesten*) (KU 5:451), or “the world a moral being would create” (R 6:5). That is the highest good as it fits the concept of a Roycean lost cause. One thing that is quite clear, however, is that the highest good is *not* to be thought of as an end *in addition to* (or over and above) other moral ends we set. The duty to set the highest good as an end “does not increase the number of morality’s duties but rather provides these with a special point of reference for the unification of all ends” (R 6:5). The highest good is rather only a way of thinking systematically about moral ends as a totality and a unity.

When Kant describes the highest good as virtue or morality combined with the happiness of which morality constitutes the worthiness to be happy, it is apt to sound as if he thinks the essential thing for which we should strive is some sort of cosmic justice. That’s how Michalson is thinking of it when he calls the highest good “cosmic symmetry.” It would be a correct way to think about it for someone who could occupy the standpoint of God; but it is a highly misleading way to look at the highest good as an end of *our* actions. We can make the highest good an end because we can promote our own moral virtue and also that of others when this coincides with their own ends; we can also act in certain ways that tend to make happiness proportionate to worthiness, as by refusing to

seek our own happiness or that of others by immoral means. But we have no duty to promote the moral perfection of others as such, and our only duty regarding the happiness of others is always to include their permissible happiness among our ends and never to make anyone's *unhappiness* an end for its own sake (MS 6:385–388). It is *never* our business to *deprive* anyone – not ourselves and still less anyone else – of some happiness just because we might think they are unworthy of it. Kantian ethics even strongly condemns any such thing. These limits on the ways in which we can pursue the highest good reinforce the thought that although our dutiful action contributes to the highest good as its final end, the possibility of highest good depends on a causal relationship *in the world* between worthiness and happiness – a relation that could exist only through the agency of a wise, powerful, and good Deity.

The question remains, however, Why should we think of the totality of our moral ends as united in this particular way, as combining two different goods, one of them conditional upon the other? And why is our pursuit of this unified end regarded as a moral *duty* bound up essentially with (and even in a way identical with) all our other duties, rather than being merely a discretionary end over and above the ends it is our duty to set? Each of the three Critiques gives a different answer to these questions, which is why we should consider the answers separately.

2.5 THE MORAL ARGUMENT IN THE *CRITIQUE OF PURE REASON*

In 1781, when Kant at last laid the foundation for his critical philosophy by publishing the first edition of the *Critique of Pure Reason*, he had still not formulated his mature moral philosophy as we find it in the *Groundwork* and later. One consequence is that his presentation of the moral argument for God, immortality, and a moral world differs significantly from later presentations. Of course Kant's 1781 position does already include several theses we find in his later ethics as well. First, Kant treats the moral law as a *law of freedom*, determining *what ought* to happen, as distinct from laws of nature, that determine what does happen (KrV A802/B830, A806/B834, cf. G 4:463). Second, the moral law commands *absolutely* – or as he will later say, *categorically*, that is, not conditionally on any end we have already set independently (KrV A807/B835, cf. G 4:414–420). Morality is thus distinct from prudence or pragmatic reason, which tells us only what we must do *in order to be happy* (KrV A805/B834, cf. G 4:416). Third, Kant also holds that moral conduct is conduct through which we become *worthy* of happiness (KrV A806/B834, A808–809/B836–837, cf. G 4:393).

In 1781, however, Kant's views about the moral incentive – the ground or reason why we should obey the absolute (or categorical) imperative of morality – are significantly different from his views after 1785. In the *Groundwork* and later, Kant famously takes the position that pure practical reason, duty, or the moral law all by itself constitutes a sufficient rational incentive for moral action (G 4:397–399, KpV 5:20–23, 31). In the first Critique's Canon of Pure Reason,

however, Kant holds that the moral incentive must satisfy the *whole* of our rational nature and not merely the pure rational or *a priori* part of it. Consequently, the moral incentive not only depends on conduct making us worthy of happiness but must in addition include the rational *hope* that we will enjoy the happiness of which we are thereby made worthy (KrV A805/833). It is this (rational, moral) hope that requires acceptance of God and a future moral world through which happiness will be proportioned to worthiness:

Thus without a God and a world that is not now visible to us but is hoped for, the majestic ideas of morality are, to be sure objects of approbation and admiration but not incentives for resolve and realization, because they would not fulfill the whole end that is natural for every rational being and determined *a priori* and necessarily through the very same pure reason. (KrV A813/B841)

Henry Allison calls our attention to this difference between the Canon and Kant's later ethics. He thinks Kant abandoned the position of the Canon because he saw it to be both eudaimonistic and heteronomous (Allison, 2011, pp. 55–56). Allison is clearly right to say that Kant's position in the Canon does differ from the one adopted later. But in the Canon, Kant already clearly distinguishes moral reason from prudential reason and the end of morality from happiness. He includes happiness within the moral end but only conditionally, just as he does later. This means that his position is not eudaimonistic, if eudaimonism is the position that morality is the same as happiness or is derived from the end of happiness (*eudaimonia*). Allison is mistaken, therefore, when he says that Kant's position in 1781 is eudaimonistic.

In 1781 Kant does not yet employ the concepts of autonomy and heteronomy. This makes Allison's other charge – that the Canon's position is heteronomous – harder to evaluate. The inclusion of hope for happiness within the moral incentive might make the ethics of the Canon heteronomous according to Kant's later conceptions of autonomy and heteronomy, if autonomy is obedience to a law one regards as self-given solely for the sake of that law and heteronomy is obedience that is not grounded solely on this. But if autonomy is conceived as acting *solely from the incentives of our reason*, then in 1781 we would be autonomous when we are motivated by the fulfillment of our whole end as rational beings, which includes morality as the worthiness to be happy together with the hope for the happiness of which we have made ourselves worthy. If autonomy is conceived in this latter way, then the Canon's conception of the moral incentive would count as autonomous every bit as much as the post-1785 conception.

Regarding the moral incentive, there are two questions that always need to be kept distinct. One question is a question of practical reason: *What rational incentive, if any, is there for us to obey the moral law?* The second is a question of contingent, empirical human psychology: *What, if anything, is capable of motivating human beings to obey the moral law?* Some philosophers might think that the answer to the first depends on some answer to the second. They

might hold that we have a genuine reason to do as we are morally obligated to do only if there is in empirical human nature some desire or sentiment that is capable of actually getting us to do it. Kant is clearly not one of these philosophers. The meaning of the famous (Kantian) claim “Ought implies Can” is that to our awareness of any specific obligation there is joined necessarily a rational incentive that is rationally sufficient to make it practically *possible* for us to comply with the obligation *as far as motivation is concerned* (for instance, see KpV 5:30, R 6:50). The empirical question whether the rational moral incentive is sufficient actually to motivate human beings is a question only about the degree of moral virtue in human beings. We will see in Chapter 3 that Kant doubts whether, as a matter of human psychology, the moral incentive is ever (all by itself) in fact a sufficient motive for any of us; but he does not doubt that it is rationally sufficient for us to *be able* to do as we ought (G 4:407, R 6:38–39). That we may be too morally weak or corrupt to do what we ought does not alter in the least what we have reason to do. Allison seems to confuse these two questions when he writes, “Kant’s view at this time was that, while not itself a directly motivating factor, since this would destroy the moral disposition, the belief in a future life, in which happiness will be rewarded commensurately to its deservedness, provides an important source of moral encouragement, apart from which moral effort could not be sustained” (Allison, 2011, p. 55). Thus the issue for Allison is not whether we have rational incentives sufficient to ground our obedience to the moral law but whether we need additional incentives, based on our hope for happiness, to sustain ourselves in the moral struggle. Belief in God and a moral world are supposed to be aids we fall back on to provide these *additional* incentives when we lack the virtue to do our duty solely for duty’s sake. As I see it, this is *never* Kant’s argument.¹¹

If Kant’s position is the one Allison describes, then it seems to me also to lack psychological plausibility.¹² Perhaps if I *knew* that there is a God who rewards virtue and punishes vice, then that knowledge would give me empirical incentives to do right and avoid wrong. But Kant’s position is that I can have

¹¹ Using belief in God to supplement our weak moral motivation is an interpretation of Kant’s moral arguments that has been explored by Adams (1987, pp. 144–164; 1999, pp. 384–391) and Chignell (2018). That suggests an argument I regard as fundamentally different from Kant’s. That the offered psychological benefit is moral may distract us from the fact that this argument is not fundamentally different from the corrupt bribe offered you to believe that seven is not a prime number. No promised advantage or benefit, even a moral one, could ever give you *any reason whatever to assent to God’s existence*. It offers an instrumental reason for doing something *corrupt and self-deceptive* to yourself that brings about *nonrational* assent. In Kant’s terms, it grounds persuasion (*Überredung*) rather than conviction (*Überzeugung*) (A821/B849).

¹² One passage in Kant’s *Lectures on the Philosophical Doctrine of Religion* suggests that moral faith gives us “more incentives” or “stronger” incentives to do our duty (VpR 28:1002–1003). I read this passage as anticipating his later view and yet also showing hesitation between the view of the Canon and the new conception of moral incentives Kant came to adopt after 1785. This passage is not consistent with the Canon or even with most of what Kant says in the *Lectures*.

no such knowledge but at most only Belief based on practical grounds. How could my moral effort be strengthened by a belief (or even Belief) in God that I know to be theoretically groundless, which I allegedly adopt only in order to influence my own motivations? Such a belief would have to involve a kind of dishonest pretense and moreover one I would have every reason to see through. If I am tempted to vice by the thought that there is no evidence that my moral conduct will be rewarded, then I am not likely to choose a better path by trying to make myself believe (irrationally) that it will be rewarded, since my temptation arises precisely from my all too vivid awareness that such a belief is evidentially baseless.

The basic issue in the Canon is not psychologically sustaining effort on behalf of morality. It is instead the nature of the *rational incentive* we have for moral conduct. The question is, What is that incentive? Kant's position in the Canon seems to be that we are aware of an unconditional *obligation* to obey the moral law, grounded on practical reason alone: "So far as practical reason has a right to lead us, we will not hold actions to be obligatory because they are God's commands, but rather regard them as divine commands because we are internally obligated to them" (A819–B847). But it is a separate question what *rational incentive* we have to do what we know we are obligated to do. The Canon takes the position that this incentive must "fulfill the whole end that is natural for every rational being." Therefore, it must include not only the rational part of our nature but also the part that reflects our finitude or sensibility and involves our desire for happiness. The two of course do not have the same rational status: as included in our moral incentive, our hope for happiness must be conditional on our having made ourselves worthy of happiness. But the Canon's doctrine is that this hope plays an indispensable role in the moral incentive. Without it, morality would offer us "to be sure objects of approbation and admiration but not incentives for resolve and realization" (KrV A813/B841). We see here that it is not a question of whether we need psychological encouragement to do what we have a rational incentive to do but whether we have the rational incentive in the first place to do what we know we have an obligation to do.

Kant's position in the Canon is closely related to his claim in the *Lectures on the Philosophical Doctrine of Religion* that the moral argument for Belief in God and a moral world takes the form of an argument *ad absurdum practicum*: "An *absurdum logicum* is an absurdity in judgments; but there is an *absurdum practicum* when it is shown that anyone who denies this or that would have to be a scoundrel [Bösewicht]" (VpR 28:1083). Or as he also puts it: "Hence without God I would have to be either a visionary [Phantast] or a scoundrel . . . I would have to cease to be a rational human being" (VpR 28:1072, cf. 28:1010).

I think we may understand Kant's argument in these passages as follows: If I were to deny God and a moral world, then (according to the doctrines of the Canon) I would lack a *rationally sufficient* incentive to do what I already know I am obligated to do. Morality would then still be an object of obligation and also approbation and admiration; conversely, immorality would be an object of

blame, disapproval, and contempt. I would admire the morally good person as virtuous and condemn the immoral person as a scoundrel. But if I were to deny God and a moral world, I would nevertheless lack a rationally sufficient incentive to *practice* morality or avoid immorality. If I then did behave morally, I would be a “visionary” (*Phantast*) – that is, a person I would have to regard as admirable but practically irrational. Conversely, if I did only that for which I have a sufficient rational incentive, I would violate my obligations, incur my own disapproval, and consider myself a “scoundrel.” Either way, I would be involved in an *absurdum practicum*, since I would act either morally but without a rational incentive or else rationally but in a way I must morally condemn. In either case, “I would have to cease to be a rational human being.” Notice that this argument in no way depends on the psychological strength of whatever might in fact motivate me (or fail to motivate me) to obey or disobey the moral law. For all this argument shows, a rational incentive to do my duty might not suffice to move me psychologically, and I might have psychologically sufficient nonrational motives to do my duty even if I had no rational incentives to do it. The “practical absurdity” arises solely from considering the rational incentives that apply to moral conduct, not their psychological strength in me.

There is even something attractive, compared to Kant’s later views, about the Canon’s idea that the moral incentive must encompass the whole ends of our rational nature and must not leave the hope for satisfaction of our finite needs entirely unaddressed. We should therefore wonder why Kant ever abandoned this position and later moved to the apparently more austere view that the moral incentive is that of duty or pure practical reason alone. Allison’s attempt to explain Kant’s change of mind makes sense only if we assume his later conception of the moral incentive as consisting in the moral law or duty alone. But Kant did not accept that in 1781, and so Allison’s explanation fails. If we think there is a problem internal to the Canon’s position, we need a different explanation.

Here is my suggestion: Kant’s position threatens us with rational (not merely psychological) instability in our attitudes. For it presents us with the thought that if we did not believe in the existence of God or hope for a moral world, then we would have reason to approve and admire moral conduct but no sufficient incentive to resolve on it and practice it. Our agent-attitudes, whether we have rational incentives to do as we ought, might then seem strangely alienated from our appraiser-attitudes concerning what we approve and admire. If we might have no incentive to practice obedience to the moral law, then why should we continue to admire and approve such obedience?¹³ The conflict might even

¹³ This seems to be the problem for Polus in Plato’s *Gorgias*. According to him, those who suffer injustice are worse off (less happy = *εὐδαιμον*) than those who do injustice, and those who do injustice are better off (or happier); but those who do injustice are more shameful (= *αἰσχροὺν*) and less admirable (= *καλὸν*) than those who suffer injustice (Plato, *Gorgias*, 475). Socrates argues that Polus’s appraiser-valuations of people as admirable or shameful are inconsistent with his agent-valuations of how we should act – what he thinks we should do or seek.

make us begin to look hypocritical: Do we approve of conduct in others that we would like them to direct toward us, while regarding ourselves as having no incentive to practice that same conduct toward them?

In the Canon, of course, Kant argues that this problem should be resolved by the moral argument for Belief in God and hope for a moral world. If I have moral Belief or faith, and hope for the happiness of which I have made myself worthy, then I *do* have a sufficient incentive to practice and realize the moral conduct that the moral law gives me reason to approve and admire. This is only another way of saying that moral faith alone enables me to avoid an *absurdum practicum*. It also means that the third of Kant's three famous questions – "What may I hope?" – has a far greater significance in the Canon of the first Critique than it does in Kant's later philosophy. For in 1781 it is only that hope that provides a sufficient rational incentive for moral conduct.

Kant reiterates this position regarding the moral incentive in 1783 in his *Lectures on the Philosophical Doctrine of Religion*: "The possibility of the concept of God is supported by morality, since otherwise morality would have no incentives" (VpR 28:1010, cf. 28:1065, 1073). Moreover, it is even as if he were eager to respond to Allison's charge that the first Critique position commits him to eudaimonism. For Kant there argues repeatedly that we can *avoid* a corrupting eudaimonism only by making our assent to the existence of God a form of *faith* rather than knowledge. For only a morally grounded faith in God can guarantee that the agent gives priority to the resolve to be worthy of happiness over the hope for that happiness. If the moral agent *knew* there is a God, he "would represent God as himself a rewarder or avenger; this image would force itself involuntarily on his soul, and his hope for reward and fear of punishment would take the place of moral motives; the human being *would be virtuous from sensible impulses*" (VpR 28:1012, 1073–1074, 1081, 1084). Only a morally motivated *hope* for happiness, based on rational *faith* rather than prudential knowledge, gives us a rational moral incentive *without* committing us to eudaimonism.

Whether or not we find Kant's position in the Canon attractive, it is clearly self-consistent. It avoids both eudaimonism and heteronomy. It may even affirm autonomy more successfully than Kant's later view if we think of ourselves as more autonomous when our action is grounded on *our entire nature* both as rational beings capable of recognizing moral obligation and as finite or sensible beings who rationally desire and hope for the happiness of which their conduct has made them worthy. I therefore do not think we can say for sure why Kant abandoned the position of the Canon. Of course he may simply have concluded *on independent grounds* that pure reason alone (duty or the moral law simply by itself), excluding all other considerations – or at least abstracting from them – is the only genuine rational incentive. That by itself would require him to modify the moral argument of the Canon. We will see in Chapters 3 and 4

that fundamental to Kant's moral philosophy after 1785 is his commitment to the crucial moral importance of a uniquely rational moral incentive that is independent of all other incentives. It is also crucial to the entire argument of the *Religion*: to the thesis of radical evil in human nature and to the change of heart as the proper response to it. Perhaps related to this worry, Kant may have decided that moral faith should not play a crucial role in the moral incentive, as the Canon requires. He may have thought that even hardened unbelievers *should* have a sufficient incentive to practice the moral conduct that, as rational beings, they cannot but approve and admire.

2.6 THE MORAL ARGUMENT IN THE *CRITIQUE OF PRACTICAL REASON*

Whatever his reasons for making the change, after 1785 Kant does in fact modify the position of the first Critique. He no longer treats the practical argument for God and a moral world as the way of avoiding the threat of instability faced by his more inclusive conception of the moral incentive. And he no longer regards our hope for the happiness of which we have made ourselves worthy as essential to our having an incentive to obey the moral law. In Kant's second and third Critiques, following the lead of the *Groundwork*, he takes the position that pure reason can be practical of itself and that duty alone ought to be (even if for us imperfectly rational beings, it never is) a sufficient incentive for obedience to the moral law. Thus our incentive for obeying the moral law is no longer argued to be dependent on the moral argument for God's existence.

Schopenhauer mounts the following criticism of Kant's moral argument: "We do not find Kant's doctrine of virtue so pure; or rather the presentation falls far short of the spirit and has in fact lapsed into inconsistency. In his *highest good* . . . we find virtue wedded to happiness. Yet the ought, originally so unconditioned, does postulate afterwards a condition for itself, really in order to be rid of the inner contradiction, burdened with which it cannot live. Now supreme happiness in the highest good should not really be the motive for virtue; yet it is there like a secret article, the presence of which makes all the rest into a sham contract. It is not really the reward of virtue, but yet it is a voluntary gift for which virtue, after work has been done, stealthily [*listig*] holds its hand open" (Schopenhauer, 1958, vol. 1, p. 524). In this passage Schopenhauer cites the moral argument from the second Critique. But his charge of inconsistency makes sense only if we think of the instability of Kant's position on the moral incentive found in the first Critique – or, better yet, if we ascribe to Kant the position on the moral incentive he took after 1785 but then try to combine this with the conception of the moral incentive (which had to include both morality and the hope for happiness) that characterized the first Critique. There really would be an inconsistency in *that* position; Schopenhauer would be entirely on target. Kant avoids the inconsistency by adopting quite different (and incompatible) accounts of the

moral incentive before and after 1785. Each is self-consistent, but they cannot be self-consistently combined.

This change in Kant's position is accompanied by another: in the second Critique the realization of the highest good for each of us is located no longer in a future life but in a natural order governed by divine providence.¹⁴ In the second Critique, immortality of the soul is seen as a postulate of practical reason necessary for complete actualization only of the first or supreme component of the highest good (morality or virtue) (KpV 5:122–124). We can conceive of complete realization of the first component only through endless progress toward holiness, represented as taking place in a future life. But the moral argument for God's existence no longer locates the happiness of which we have made ourselves worthy in a future life. It postulates God's existence in order to justify assent to the claim that happiness is apportioned to worthiness in this life. The moral argument for God's existence in the second Critique and afterward is therefore simpler in structure. Practical reason, namely, can justify theoretical assent based solely on the choice to set an end and the rational requirement that when I set an end, I must not only not be certain that it is impossible but must also regard myself as rationally required to assent to the proposition that the end must be possible of attainment.

The second Critique argument may be summarized as follows:

1. I have an obligation to will the highest good as an end. (In the *Critique of Practical Reason*, this follows from the rational principle of seeking unconditioned totality in our concept of the object of pure practical reason: KpV 5:108.)
2. To will something as an end presupposes that it is possible – and that I have the means to contribute to its actuality (G 4:394, 437). (Based on the concept of willing an end.)
3. Therefore, someone who wills the highest good as an end must assent to the conditions of its possibility (KpV 5:113–114, 119). (Follows from 2.)
4. The highest good is possible only if there is a God who apportions happiness to worthiness. (Postulate: immediately known condition for achieving a proposed end.)
5. Therefore, if I fulfill my obligation to will the highest good as an end, I must assent to the existence of God. (Follows from 1, 3, 4.)

¹⁴ Of course Kant obviously does not abandon immortality of the soul in the second Critique. It is even the *first* Postulate of Practical Reason. But now it is grounded solely on our need to think of a future life in order to make possible an endless progression toward moral perfection, not so that happiness may be apportioned to worthiness (KpV 5:122–124). In the *Religion*, however, moral progress occurs in this life and apparently no longer requires an endless future life (R 6:67–71). In the third Critique, the highest good (“the human being under moral laws”) unifies moral teleology with natural teleology, implying that the apportionment takes place entirely within the realm of nature (KU 5:447–453). Kant still mentions immortality (KU 5:470, 473–474), but it is no longer clear (at least to me) why. Compare Wood (1970, 2009), pp. 181–182.

Comments on (1) and (4) are called for. As to (1), Kant is clear, as we saw, that making the highest good our end does not add to our duties but only provides them with a systematic focus. Someone might therefore omit to make the highest good their end without being *morally* at fault, even if Kant might criticize them *theoretically* from the standpoint of rational systematicity. This is one reason Kant would insist that Belief cannot be morally commanded (KpV 5:144) and why failure to accept the moral argument is not by itself a violation of duty.

Regarding (4), the *Critique of Practical Reason* refers to the existence of God as a “postulate of pure practical reason” (KpV 5:122–134, cf. KrV A634/B662, VpR 28:1083, O 8:141). In Kant’s logic, practical propositions – those representing some action as possible according to reason – are divided into “problems” and “postulates.” A *problem* (*Aufgabe*, *problema*) is a *demonstrable* proposition expressing an action whose manner of execution is not immediately certain (VL 9:112).¹⁵ A *postulate*, by contrast, is an *immediately certain* and *indemonstrable* proposition determining a possible action whose manner of execution is also immediately certain (VL 9:112). In the Dialectic of the second Critique, the action in question is that of willing the highest good as an end. It is immediately certain both that and how we may do this – namely, by obeying the moral law and willing the other ends (duties of virtue) it commands us to will. But it is indemonstrable that this end is possible, since it depends on the existence of God, which is indemonstrable by theoretical reason. The postulate of pure practical reason thus represents to us as possible the condition for performing the action of willing the highest good as an end. And because this action is morally required, the postulate becomes immediately certain for practical reason, even if it remains indemonstrable and uncertain for theoretical reason.

It remains unclear, however, whether the immediacy and indemonstrability is supposed to attach also to the connection between willing the highest good as an end and the postulation of the condition of the possibility of rationally willing this; for it does not seem immediately certain, or in no need of demonstration, that God’s existence is such a condition. On the contrary, it looks as if Kant presents an *argument* that the highest good is possible *in the world* only on the condition that there exists a highest *original* good, having the knowledge, power, and will to make it possible (KpV 5:124–125). The classification of the moral argument for God’s existence in the second Critique as a “postulate” implies it is *immediately certain* that God’s existence is required for the possibility of the highest good. If a distinct demonstrative

¹⁵ In the Dialectic of the *Critique of Pure Reason*, the cosmological ideas (the beginning of the world, its limit in space, simple substance, free causality, necessary existence) are given as problems (*aufgegeben*) by pure reason (A498/B526, cf. A335/B392). That is, reason demands that we seek their objects; how we do so is not immediately certain, nor is it immediately certain whether we can perform the action in question; but it is *demonstrable* how we search for these objects – namely, through regulative (not constitutive) principles of reason (A508–515/B536–543, cf. A624ff/B670ff).

argument were needed to show that the highest good is possible only if there exists a God in the sense of a traditional *ens realissimum*, then it would no longer be accurate to describe the moral argument for God's existence in the second Critique as a *postulate*. But Kant's use of this term thus once again reveals the depth of his commitment to the traditional scholastic-rationalist concept of God.

Some scholars have doubted that anything like this postulate, whether immediately known, demonstrated, or simply granted for the sake of argument, is needed for our pursuit of the highest good. Andrews Reath has distinguished a *secular* conception of the highest good, in which it might be seen as attainable by collective human striving, from a *religious* conception, which is thought to depend on God's agency (Reath, 1988, pp. 593–619). Reath argues that if attainment of the highest good is within our power (the secular conception), then we do not need a postulate of the divine in order to pursue it, but if it is not (the religious conception), then there cannot be any duty for us to set it as an end, since no action of ours could achieve it. Reath's argument is similar to Lewis White Beck's argument that our duties could not include pursuit of the highest good if its achievement is something for which God's action would be required: "For suppose that I do all within my power – which is all any moral decree can demand of me – to promote the highest good, what am I to do? Simply act out of respect for the law, which I already knew. I can do absolutely nothing else toward apportioning happiness in accordance with desert – that is the task of the moral governor of the universe, not of a laborer in the vineyard" (Beck, 1960, pp. 244–245).

Despite its secularist appeal, however, this line of thinking won't hold water. The response is that in order to set the highest good as an end, I do not need to suppose that I alone, or even the human race alone, can achieve it. It is enough that I can regard my actions as making some contribution toward it. The same is true of many ends that are humbler and more short term. To do my part toward achieving some political goal (say, passing some piece of legislation), I need not think the goal is achievable by me alone, but only that my actions can contribute to it. In thinking of the highest good as my end, I do think of my dutiful actions as making a contribution toward it; I do not know how these contributions relate to the contributions of other human beings or how they relate to God's providential design for the universe.

Thus there is no ground for Reath's distinction between a merely "secular" conception of the highest good, for which merely human effort might suffice, and the larger ("religious") conception, for which we cannot reasonably suppose it does. Beck's metaphor of the laborer in the moral vineyard betrays the basic misconception here. The Kantian moral agent, as "laborer in the moral vineyard," is not a mere slave or tool used by the moral governor of the universe for the achievement of the governor's ends. On the contrary, the moral agent has a duty to share with all other rational agents (including God) the very same complete end. In promoting the highest good,

I must see myself as working along with other moral agents and with God as well. I am not able to know precisely what my contribution is; but I do rationally require the Belief that the end is possible of attainment and also that the conditions for this exist – including, of course, the crucial contribution of God. I can draw no sharp distinction between my contribution and God's not only because I lack the cognitive capacity to draw it but also because drawing it so sharply as to leave everything to the Deity, as Beck suggests we do, would amount to my refusing to regard the highest good as *my* end at all.

2.7 THE MORAL ARGUMENT IN THE *CRITIQUE OF THE POWER OF JUDGMENT*

In the third Critique, the moral argument seems again to be essentially the same as that in the second Critique – the five-step argument summarized above. But it is given a different and larger context, namely, the attempt to systematize all teleology – combining the purposiveness of nature with that of reason or morality – into the concept of an ultimate end (*letzter Zweck*) for the entire world. Kant's strategy for unifying ends seems to be quite close to that used by Aristotle at the very beginning of the *Nicomachean Ethics* (1094a–b). All ends are combined by subordinating the actions, arts, or teleological systems they involve to other actions, arts, or teleological systems, resulting in a single ultimate end for all of them. In the *Nicomachean Ethics*, the art of bridle-making is subordinated to that of horsemanship and horsemanship to military strategy, and, in a similar way, the ends of all human striving are systematized under the end of happiness (*eudaimonia*) – especially the larger or more complete happiness of a society or a polis. Kant argues that the end under which we may systematize all the ends of nature is the human being because of all natural beings, the human being is the only one capable of consciously willing ends and organizing ends into a system. This necessarily implies that the human being is the ultimate end of nature (KU 5:426–431).

Kant then asks what more specifically it is about the human being that could be such an end. He rejects the suggestion that human *happiness* constitutes such an end. For nature has not arranged things so that happiness is even among nature's ends, much less its ultimate one. The human pursuit of happiness, which never reaches complete success, is rather nature's means for developing human capacities – collectively and in history. Human beings constitute the ultimate end of nature, therefore, only through the culture (*Kultur*) through which they develop and exercise these capacities according to reason. It is the *vocation* (*Bestimmung*) of human beings to make themselves into the ultimate end of nature through this culture (KU 5:431).

Human culture takes two basic forms. The first is the culture of *skill*, through which human beings develop the capacity to pursue ends set at their discretion;

the second is the culture of *discipline*, through which human beings constrain themselves to will ends determined by reason (KU 5:432–433). The former culture is clearly subordinated to the latter. Therefore, the human being becomes the ultimate end of nature only through the culture of discipline. Reason's highest ends are those of morality; the highest of these is their sum total, the highest good in the world (*Weltbeste*) (KU 5:451), which Kant now renames as “the human being under moral laws” (KU 5:445). This formulation includes human beings who do not act in accordance with moral laws but enjoy only a happiness proportional to the morality of their conduct (KU 5:448–449 n). The highest good, so understood, is the only end that may be regarded as a *final end* (*Endzweck*) – that end which needs no other end as a condition of its possibility (KU 5:434). For it includes all the ends that can be rightly set by reason. It is therefore only through willing the highest good that humanity can fulfill its vocation. This vocation is not only *to be the ultimate end* of nature but to be this end precisely *by setting a final end* for the whole of creation (KU 5:435, 437, 443).

In the third Critique, the moral argument is further developed through what Kant calls an “ethicotheology” – a conception of God as the being whose volition enables us to understand nature and our own place in it through the concept of a teleological system (KU §86). Clearly this is not Michalson's “shrunk” God, a lesser deity in comparison to the traditional metaphysical concept of an *ens realissimum*; on the contrary, it is a version of that very concept – one which symbolizes God according to what is suitable to our moral vocation. If we are religious, we look on God as commanding us to fulfill our vocation in relation to nature. He is our judge under moral laws for the manner in which we conduct ourselves in relation to the teleological system of nature (KU 5:446). Our need to accept (*annehmen*) a moral cause and ruler of the world is justified by our need to think of our vocation through feelings of *thankfulness* (for nature's purposive harmony), *obedience* (to the moral law of reason), and *humiliation* (self-reproach at our moral failings and deserved *chastisement* when we fall short of our vocation of being the ultimate end of nature and giving it a final end) (KU 5:446–447).

Kant's thesis that only the human being can be the ultimate end of nature, and give the whole of creation a final end, is the diametrical opposite of the truly horrid idea that nature exists only in order that human beings may impose their arbitrary, exploitative, and short-sighted will upon it. On the contrary, the vocation of human beings is to preserve and promote that harmonious system of ends that already exists in nature. In order to do this, we must curb our propensity to make our own inclinations and happiness our chief object of concern and instead constrain our wills to make ourselves into that by which nature itself constitutes a harmonious system. Thus it is a basic duty to ourselves, as the ultimate end of nature, to preserve the beauty of nature and not to uproot our natural disposition to appreciate and promote it (MS 6:443). The *Critique of the Power of Judgment* continues to speak of “assent” or

“acceptance from a practical point of view” as the propositional attitude produced by the moral argument (KU 5:452–453). In the third Critique, however, Kant also goes beyond this, depicting the emotional engagement of a morally disposed man (he refers here, with very dubious plausibility, to Spinoza) who is intellectually disposed to reject the existence of God but might be converted to belief (or at least to Belief as practical acceptance) by attempting to align his theoretical doubt of God with his moral intentions (KU 5:452–453). Kant even goes so far as to describe rational moral faith in God as a habit or aptitude (*habitus*) and as a “way of thinking” (*Denkungsart*) that is disposed to affirm God’s existence on practical grounds, even though such belief is “inaccessible for theoretical cognition” (KU 5:471–472). He suggests this faith may be mixed with doubt (KU 5:472–473).

2.8 THE KIND OF ASSENT JUSTIFIED BY THE MORAL ARGUMENT

In all three Critiques, the primary attitude of assent involved in practical commitment is “faith” or “belief” (*Glaube*). In all three Kant also explains what he means by *Glaube* by appending to his statement of the moral argument a set of reflections on the kind of *assent* (or holding-for-true, *Fürwahrhalten*) it justifies (A820–832/B848–859, KpV 5:143–146, KU §§90–91, 5:461–475). Again, the kind of assent he is trying to explain is that for which, following Chignell, I am using the term “Belief” (capitalized).

Before we look at these discussions of what Kant means by practical faith or Belief, it is worth pausing to consider first a different propositional attitude – namely *hope*. Both Belief or faith (at least in Kant’s technical sense) and hope are, however, clearly aims of Kant’s practical arguments. Both, moreover, seem to be, or at least to involve, kinds of assent (*Fürwahrhalten*) for practical purposes. Hope is often directed to the possibility even more than to the actuality of that for which we hope.¹⁶ Of course, we hope that what we hope for is, was, or will be actual and not merely possible. But hoping is always related first to possibility and only through it to actuality. We can’t rationally hope for what we know is certainly the case, just as we can’t hope for what we know is not the case, especially for what we know is not possible. For hope, possibility is always on the necessary itinerary toward actuality. The focus on possibility applies especially when we hope regarding future contingent events: hoping your team or party will win the game or the election is focused on the possibility of their winning and is much more closely connected to your practical efforts than believing they will win. At least in the ordinary sense of

¹⁶ See Kant’s letter to F. C. Stäudlin of May 4, 1793 (Ak 11:429, cf. VL 9:25). The importance of hope to Kantian practical faith has been especially stressed by Chignell (2015). Chignell’s conclusion is that on Kantian practical grounds, there is more room for *hope* than *belief*. In some respects, he is certainly right. See next note and also Chapters 5 and 6 below.

“belief,” I can believe the opposing team or party will win, but I certainly don’t hope they will.

When it comes to matters with which Kant is principally occupied in the *Religion*, Kant does tend to speak of *hoping* as well as, or sometimes even instead of, believing. Kant sometimes claims that Belief or faith (*Glaube*) is a concern of morality but not of religion; the properly religious attitude is instead one of *hope*. We cannot presume to believe (or even Believe) that we will become well pleasing to God (R 6:76). Our properly humble attitude toward God’s forgiving acceptance of us is instead one of *hope*, not confident belief or even Belief. Stressing the role of hope specifically in religion also makes more sense of Kant’s claim that the “minimum of theology” required for religion is not that God exists but that God is possible. Those who do not believe in God’s existence, as long as they think it is possible, might still hope that God exists, and their hope might be enough to make it rational to set the highest good as an end and also to regard their duties as God’s commands.¹⁷ Our concern in this and the following section of this chapter, however, will be not with the kind of assent involved in hope but with the kind involved in practical Belief or faith, since in the moral argument for God’s existence in all three Critiques, Kant focuses on Belief or faith (*Glaube*).

Assent (*Fürwahrhalten*) is for Kant the basic affirmative propositional attitude (as it is for Descartes and for Locke). Assent can be mere *persuasion* (*Überredung*) if it has no rational grounds but only subjective (psychological) causes that explain the subject’s assent. (If I were offered a million dollars for believing that 7 is not a prime number and successfully underwent indoctrination that gained me this money, my belief would be a case of persuasion. So would be the belief of some that global warming is a hoax perpetrated by the Chinese, when this plainly false belief is due to the way that holding it serves their economic or political interests.) Alternatively, assent can be *conviction* (*Überzeugung*) if based on genuine rational grounds of some kind (KrV A820–821/B848–849, KU 5:461). Because it involves some kind of rational justification, conviction licenses *assertion* (to others) of that to which we assent, but mere persuasion does not. If we realize we are merely persuaded of something, we might confess to others that we assent to it, but we have no business asserting it, since that carries with it the implication that others should assent to it too. A fuller account of this distinction, and the distinctions drawn in the next few paragraphs, can be found in Chignell (2007).

Conviction comes in three grades or degrees, depending on the kind and strength of the rational grounds for the assent. *Knowledge* (*Wissen*) is assent

¹⁷ In *Kant’s Moral Religion* (Wood, 1970), chap. 6, pp. 232–248, I spoke of a “postulate of divine grace,” parallel to the other practical postulates at KpV 5:122–134. I also spoke of rational *trust* in divine grace and *faith* in it. I did not mention hope, but I certainly should have. In part two, section one, of the *Religion*, where he discusses grace, divine aid, and justification (R 6:60–78), Kant in fact speaks regularly of all three: *Glauben*, *Vertrauen*, and *Hoffnung*.

based on grounds sufficient to justify the assent and that are objective, both in the sense that they are the kinds of grounds that are valid universally for all subjects and in the sense that they relate to the objects the judgment is about (KrV A822/B850, cf. KU 5:467). These grounds may be either a priori proof or empirical grounds, including arguments from analogy (KU 5:463–465). Belief (capitalized, i.e., *Glaube*) is based on grounds that are subjective in the sense that they are based on the practical ends of the subject. These considerations are subjective even when the setting of the end is required by reason because the grounds of assent are still dependent on a choice by the subject. Above we saw how a physician's need to treat a patient might lead her to assent to a certain diagnosis; assent based in this way on a discretionary end Kant calls *pragmatic Belief* (KrV A824–825/B852–853). Kant also recognizes what he calls *doctrinal Belief*, based on theoretical aims (KrV A825/B853); he thinks that the existence of God can be assented to in this way insofar as the idea of God is used regulatively for theoretical purposes (KrV A826/B854). Here, too, the Belief is based on aims of theoretical inquiry and is therefore subjective. *Moral Belief in God* is the kind of assent produced by the moral argument. The lowest degree of conviction is *opinion* (*Meinung*) (KrV A822/B850, KU 5:465–466). The grounds for it are insufficient subjectively as well as objectively; that is, the subject is aware that they do not fully justify the assent, but opinion is still a form of conviction in cases where it can be based on a rule of evidence that might serve to justify it if only the conditions falling under the rule met standards of sufficiency. *Hypotheses* may be asserted in the form of opinions if they meet this condition (KU 5:466).

As I have already said, what Kant calls “belief” (*Glaube*) is not precisely what we normally mean by the word “belief” – or what would be meant by *Glaube* in everyday German. When we say we *believe* something, we usually mean that our disposition to assert it simply expresses the way we see the world. It forms part of the background against which we proceed in both acting and inquiring. This is why we experience belief passively and why we do not normally suppose people can change their beliefs by choice. For the same reason, it runs counter to the ordinary meaning of “belief” to say that we could believe things only for certain purposes or only in certain respects. Belief is that to which we assent by default, so to speak, and without any additions or qualifications. It is precisely *not* that to which we assent only for certain purposes or in certain respects. If we say that an assent is “absolute” when it is not done relative to any purpose or in any limited respect, then “belief” (lower case, in the ordinary sense of the word), is always *absolute* assent. It is that to which we are disposed to assent by default – not relative to any respect or purpose. Beliefs lose their background or absolute status, and become the focus of attention, only when we call them into question. We need then somehow to prop them up, or else we experience them as beliefs mixed with doubt. Then it becomes problematic whether we any longer fully or truly *believe* what we are saying we believe.

Kantian Belief, however, is supposed to be consciously an *assent only for practical purposes*, assent *in a practical respect* and relative to certain *subjective*

conditions. “Belief” in Kant’s technical sense of the word (“Belief” capitalized) is *never* absolute because it is always assent in a practical respect or for a practical purpose. In the case of pragmatic Belief, the doctor might say, “For the purpose of treating this patient, I proceed by assenting to such-and-such diagnosis.” If asked, using in the ordinary – background, default, or absolute – sense of “belief,” whether she *believes* the patient has that disease, she might very well say, “I don’t know what disease the patient has, and I don’t honestly have any *belief* about that. If I were a detached observer (like my esteemed colleague here, who has no duty or interest in treating the patient), I would have to suspend judgment about what the ailment is. It is solely on pragmatic grounds, and solely for the purpose of treating the patient, that I must rationally assent to this diagnosis.”

Analogously, Kantian doctrinal beliefs are acts of assent only for the purposes of inquiry. That is, we make use of assent for the purpose of framing our hypotheses or directing our questions. If we are critical, we do not let the assent exceed the limits of these purposes. Moral Belief, too, as Kant repeatedly insists, is held only “subjectively,” and “in a practical respect” or “for practical purposes” (KrV A823–829/B851–857, O 8:139–142, KpV 5:142–146, KU 5:467–473, VpR 28:1082–1084). Regarding moral Belief in God, we assent to the existence of God for the purpose of resolving the tension between our willing the highest good as an end and our theoretical ignorance regarding the possibility of this end. Kant sometimes refers to this moral Belief as “a felt need of reason” (O 8:139–140, KpV 5:142–143). We need to assent to God’s existence in order to avoid an incoherence in setting the highest good as an end. The need is relative to our choice set this end; in that sense it remains “subjective” even when the choice is one required by an objective moral duty. Accordingly, Kant says that the moral argument does not justify saying, “*It is certain that there is a God,*” but only that “*I am morally certain*” (KrV A829/B857). In all cases of *Glaube*, in Kant’s sense of the term, the assent is at least as much about my practical situation as it is about the subject matter of that to which I assent. It is therefore misleading to call Kant’s moral Belief “belief” at all, if we intend to use the word in the ordinary sense. This is important in deciding how far Kant’s practical faith or Belief is consistent with evidentialism – the issue we will be discussing in the final section of this chapter.

Another term Kant often uses to express Belief – that is, assent for practical purposes – is “assume” or “accept” (*annehmen*) (A811/B839, KpV 5:120–121, KU 5:453, 471, 480). This does fit our ordinary uses of these words. We can for certain purposes assume or accept (or in German, *annehmen*) even what we clearly *do not* believe. Perhaps the clearest case is this: we can assume or accept something as a premise at the beginning of an indirect or apagogic proof that the negation of this premise is true. Normally, we positively *disbelieve* that to which we assent in this case because we expect to reduce to absurdity what we accepted or assumed, precisely so as to prove that it is false. Of course Kant does not think we positively *disbelieve* the

conclusion of the moral argument. But if our attitude is Belief – assent only for practical purposes – then it is not the kind of default or absolute acceptance for which we ordinarily use the word “belief.”

Perhaps what many religious people want from their faith is an absolute assent. They want a *belief* that protects them from the anxiety and torment of religious doubt and makes them feel safe and secure in a human condition characterized by anxiety about what ultimately concerns them and threatens their lives with meaninglessness and despair. If this is what they want, then we must openly confess that the Belief supported by Kant’s moral arguments is unable to give it to them. At times, Kant seems to be trying to convince his audience (perhaps also himself) that the moral arguments do provide this sort of security and comfort (KrV A828–829/856–857, cf. VpR 28:1010–1012, 1082–1084, 1126). This is consistent with his position, however, only if the comfort in question consists in the avoidance of total despair over our end which deprives us of the dismal comfort of resignation, since if we despair we can comfort ourselves with the thought that we need no longer care about that which is utterly beyond hope. Kant’s moral arguments do intend to rescue us from this wretched kind of consolation. But they simply cannot deliver the comforting confidence of unquestioned certainty, if that’s what people want religious consolation to be. We should worry, however, about whether it is consistent with a genuinely religious attitude even to want the consolation of comfortable complacency. Religion should be more honest than that.

Kant sometimes explicitly regards what the moral arguments produce as a practically motivated and emotionally charged acceptance that is less than certain and also compatible with serious theoretical doubt (see KU 5:452–453). When religious faith is called into question, or when it is “tested” by events or troubling thoughts that tempt the believer to lose faith, then it ceases to be “belief” in an unproblematic sense, and it once again becomes “belief” in the ordinary sense of that word only to the extent that these temptations can be resisted.

Perhaps this troubled sense of “belief” or *Glaube* is also part of the ordinary sense these words can bear in certain limited contexts, where the issue is precisely one of attempting to preserve a belief (in the passive, absolute, and background sense) even in the presence of troubling doubts that threaten it *as a belief*. We then call it “belief” precisely with the aim of voluntarily assenting to it for the purpose of dragging it back from the abyss of doubt and retaining it as a belief in the ordinary sense. This might also be true of our Belief in God or our Belief in some cause to which we are devoted. It is in this spirit of treating Belief as a kind of belief-under-threat that Kant insists that moral faith in God is incompatible with *unbelief* (*Unglaube*), since that would involve the outright denial of the real possibility of the highest good, which is incompatible with still setting it as an end. But Belief is also not to be confused with credulity (*Leichtgläubigkeit*), an absolute assent which lightheartedly ignores substantial grounds for doubt. The very lack of theoretical grounds for assent

makes Belief (assent in a practical respect) quite compatible with a “doubting faith” (*Zweifelglaube*) (KU 5:472–473). What we Believe (solely in a practical respect) is precisely that which, on theoretical grounds, we cannot absolutely *believe* but must also at the same time doubt, since the theoretical grounds for it are insufficient.

Religious people – at least the wiser among them – recognize that faith is related in complex ways to both belief and doubt (see Alston, 1996). Paul Tillich argued that faith is a form of courage in the presence of anxiety (Tillich, 1952) and that doubt is an element of faith (Tillich, 1957, pp. 18–24).¹⁸ Anne Lamott says, “The opposite of faith is not doubt, but certainty” (Lamott, 2005, pp. 256–257). Kantian Belief belongs to that religious tradition.

2.9 IS PRACTICAL FAITH ANTIEVIDENTIALIST?

If we take “belief” in what I have been calling our ordinary sense, however, its opposite is doubt (or else disbelief); and certainty is the most perfect or complete form of belief. It is belief in this ordinary sense that evidentialists have in mind when they say, in Clifford’s words, “It is wrong always, everywhere and for anyone to believe anything upon insufficient evidence” (Clifford, 1999, p. 77). Kant’s moral argument for Belief in God offers grounds for assent that are explicitly practical rather than theoretical or evidential. Kant holds the quite strict position regarding transcendent metaphysics that there can be no evidence or proof of either God’s existence or nonexistence. It might therefore look as if Kantian Belief commits us to violating Clifford’s evidentialist principle.

We have seen, however, that Kant’s Belief is not the same as “belief” in this ordinary sense. Belief (capitalized), or assent in a practical respect, is marked off from “knowledge” by the fact that its grounds are practical and subjective rather than theoretical and objective. Assent, of which Belief is a species, is also a different *kind* of attitude from “belief” in our ordinary sense. In the ordinary sense, belief is a dispositional state. It is, roughly, a standing *disposition* to assent and to assent *absolutely* – not in any particular respect or for any purpose. Beliefs in this ordinary sense are often called upon to justify doing or thinking something. They themselves can become the focal point of our thoughts or deliberations but only when they point the way to *something else* that they justify thinking or doing. As we’ve seen, when a belief in the ordinary sense becomes the focus of critical attention, it is already being put in question and is no longer unproblematically a belief. Where we attend to a belief for the

¹⁸ Robert Adams thinks of *faith* as the virtue of persisting in a belief you ought to hold even in the face of substantial grounds for doubt (Adams, 1999, pp. 373–389). “Faith” in Adams’s sense is an important virtue and by no means limited to religious contexts. It is an important component of the virtue of persistence in the face of adversity and doubt that others have called “grit” (see Morton and Paul, 2019). That evidentialism must rule out anything like grit is a common error about it, first, because, as Morton and Paul argue, evidence can be permissive and also because, as I here argue, Kantian Belief, while in tension with evidentialism, is not inconsistent with it.

purpose of putting it in question, this is the kind of case Sartre has in mind when he says, "Every belief is a belief that falls short; one never wholly believes what one believes" (Sartre, 1956, p. 69).

Kant's moral Belief in God consists in the act of assenting to the existence of God for practical purposes. Assent is an *action* – sometimes public and verbal, but the assent we are talking about here – sincere inner assent – is often tacit and to oneself alone. Unlike belief in our ordinary sense, assent is a *voluntary* action, something we can do on the spot just by *choosing* to do it for some reason or other. Descartes's *Meditations* is largely an extended and disciplined exercise in acts of assenting and withholding assent, acts that are done for the definite reasons that Descartes's meditator carefully scrutinizes as he decides whether to assent or not to assent. That is also precisely why assent can be done for certain purposes or in certain respects. In the case of Kantian acts of acceptance (*Annehmen*) for practical purposes, it is an act for which there are quite specific practical reasons – namely, in order to avoid the irrationality that would be involved in setting the highest good as an end without assenting to the conditions of its possibility. Kant also often emphasizes that moral faith or Belief is assent *only in a practical respect*. We do it as part of our action of pursuing the highest good; we must not do it as part of any acts the reasons for which are theoretical proofs or evidence for that to which we assent.

Of course, if our lives are devoted to our moral vocation, then they may be seen as a pervasive pursuit of the highest good. Acts of assent for practical purposes may then become as habitual as acts of assent that grow out of belief (in the ordinary sense). Our moral life is (or should be) continuous with our entire lives; so moral Belief, unlike pragmatic Belief, is a disposition to assent for practical purposes that can be constant and pervasive for a human being. Even doctrinal Belief may apply only sometimes, when we are engaged in inquiries that require it. This may be why Kant considers doctrinal belief "unstable" in a way that moral belief should not be (KrV A827/B855). Only moral Belief is in a position to be something sufficiently habitual that it approximates to belief in the ordinary sense.

Since Kantian Belief is never an absolute disposition to assent, we should be able to see clearly on reflection that moral Belief does not violate Clifford's evidentialist principle. The evidentialist principle does not forbid acts of assent for certain purposes. It does not condemn acts of assent for practical purposes or in a practical respect, unless these have resulted in a dispositional state of absolute assent ("belief" in what I've called the ordinary sense) in something for which the evidence is insufficient. Clifford himself speaks with approval of acts of assuming and hypothesizing – which are clearly species of assent for certain purposes – whose objects are not something believed but part of an inferential process that is indispensable for the acquisition of evidentially justified belief (Clifford, 1999, pp. 92–96). Clifford does not speak of assent for *moral* purposes, but on reflection we should be able to see that his evidentialist principle is also quite consistent with its possibility.

The important thing is that practical grounds for assent are *genuine rational grounds*, giving us genuine reasons for assent that are distinct from all theoretical grounds. These grounds consist in the incoherence of setting an end while omitting to assent to the conditions of its possibility. It is also important to see how these grounds are distinct from any rational grounds we might have for producing in ourselves an irrational disposition to assent based on some benefit or advantage that might come from having such a disposition. Such a disposition would be irrational even if the ends served by acquiring it were thoroughly rational ends or even morally admirable ends, even pursuit of the highest good.

It might seem, however, that Kantian moral Belief must involve, at least as a disposition or even as a goal, a *belief* in the ordinary sense and moreover a belief for which there is no sufficient evidence. In the *Critique of Practical Reason*, Kant presents the result of the moral argument as a *decision* to assent to the truth of the practical postulates. He denies that either belief or assent can be *commanded*, but he does describe the result of the moral arguments as “maxim of assent for moral purposes” and a “voluntary (*freiwillig*) determination of our judgment” (KpV 5:144–146). This looks very much like a resolve to adopt a disposition to assent, based on practical grounds alone and based not at all on theoretical proof or evidence. Moreover, as presented in the *Critique of the Power of Judgment*, faith (*Glaube*) is described not as an act but as a *habitus* or a way of thinking (*Denkungsart*) (KU 5:471). When we Believe – assent to the existence of God for practical purposes – the expected result of a habitual series of such acts of assent is that we acquire an aptitude (*Fertigkeit*), a way of thinking involving a settled disposition to assent. The result of this process would strongly resemble “belief” in the ordinary sense. It would be a standing disposition to assent adopted on grounds Clifford’s principle would not consider “sufficient evidence.” If evidentialism were to stand in the way of this, the danger that it might paralyze us in some such way is the main point William James is making against it in his famous essay *The Will to Believe*. Framed in these terms, James is making a valid point.

The *errors* James commits in that essay, however, are many and serious. They are usually overlooked or interpreted too charitably. To cite only two: *first*, James insists we must approve *any* belief to which we might be psychologically tempted; *second*, he considers beliefs rational when they are adopted on the grounds that they are *advantageous*. Both these kinds of corrupted judgment are precisely what Clifford means to condemn – and rightly. Violations of the evidentialist principle are probably the commonest and most harmful forms of moral laxity with which human beings are afflicted. People usually overlook this because there is normally nothing overtly cruel simply in failing to believe according to the evidence. But such failures are often complicit in precisely these more obvious forms of evil and are commonly used to license them before consciousness. Violations of Clifford’s principle are among the most insidious, as well as the most common, form of evil.

To see an overt inconsistency between Kantian moral faith and Cliffordian evidentialism, you would need to have a misconception, or at least a shallow and inadequate conception, of both. Evidentialism is an insistence on questioning and inquiry. We must not refuse to question, or refuse to be open to the evidence. That is not incompatible with being practically committed to those ends you ought to pursue. Conversely, the commitment to those ends is not incompatible with being soberly honest with yourself about the likelihood of achieving them. Honest commitment to an end should be entirely compatible with an open-eyed refusal to entertain illusions about the likelihood of success. If someone were to suggest that the strength of our commitment to the highest good ought to be proportional to our comforting certainty that there is a God who will bring it about, then it is *morally* important to deny this. For a state of comforting certainty tends, on the contrary, to weaken or destroy our commitment to pursuit of an end. Confidence is needed only to the degree that confidence stands opposed to despair, giving up on our end, resigning ourselves to its hopelessness, falling into the indifference that might be born of the thought "It is too painful to care." It is the choice of comfortable despair or indifference that saps practical commitment, in just the same way complacent certainty might do. Sober worry based on the uncertainty of success, the honest resolve not to believe disproportionately to the evidence – these are entirely consistent with Kantian moral Belief.

Kant's theoretical agnosticism about God's existence does not supply us with sufficient evidence for belief in God's existence – in the ordinary sense of "belief." But it might be sufficient to justify Belief: assent for practical purposes. It might even be sufficient to justify a habit, disposition, or "way of thinking" involving habitual acts of such assent. That would be compatible with evidentialism, as long as it did not involve a flight from questioning or an attempt to protect oneself from the doubts arising from a lack of theoretical proof or evidence. It might seem paradoxical to say that theoretical agnosticism about something could ever constitute "sufficient evidence" for assent to it. But even Clifford would say it could be sufficient for acts of assent to something as a hypothesis. And it might be entirely rational to commit yourself to an end – in Royce's terms, to loyalty to a lost but not hopeless cause – even if you were theoretically agnostic about the likelihood of its success.

Practical Belief and evidentialism might complement and even require each other. For they have a common enemy – namely, *comfortable complacency*. Both aim at *unsettling* us – whether by prodding us with hope or troubling us with doubt – so that our concern and effort are not undermined either by despair or by overconfidence. Despair can be comforting: A weak king, facing defeat, sought a kind of comfort in hopelessness: "Beshrew thee cousin, which didst lead me forth from that sweet way I was in to despair" (Shakespeare, *Richard II*, act 3, sc. 2, ln. 204). Despair was sweet for pitiful King Richard because hope demands something of us, at least of our feelings. There is equal danger in the comfortable complacency arising from confidence that your end

will inevitably be achieved even without your efforts. This is the attitude Isaiah Berlin famously attacked in his essay “Historical Inevitability” (Berlin, 2002, pp. 94–165). It was never this attitude that Karl Marx was advocating when he declared the fall of capitalism to be inevitable. As Terry Eagleton has said, “Historical determinism is a recipe for political quietism” (Eagleton, 2011, p. 46). Martin Luther King Jr. famously declared, “The arc of history is long, but it bends toward justice.” Barack Obama significantly added, “But it does not bend of itself; *we* have to bend it.”

There is a tension between Kantian moral Belief and evidentialism, since the former offers us grounds for assent that do not consist in evidence. But that tension ought to help us find the doxastic virtue lying between the twin vices of uncritical credulity and stubborn incredulity. Rational action requires continual, even habitual assent to the possibility of our end but also freedom from self-deception about the likelihood of success. The tension will produce precisely that practically committed *Zweifelglaube* whose diametrical *opposite* is comforting certainty – whether the certainty of deluded overconfidence or that of despairing resignation.

The basic insights underlying both Cliffordian evidentialism and Kantian moral faith are therefore of immediate and vital relevance to anyone committed to a meaningful cause of any kind. If people saw this clearly, they would cease both their dishonest evasions of Clifford’s ethics of belief and also realize that Kant’s moral arguments are not deserving of the neglect and contempt commonly shown them.

The Radical Evil in Human Nature

In the literature on the *Religion*, it is usually taken for granted that Kant is fully committed to the thesis that there is a radical propensity to evil in human nature. Scholars then puzzle over how Kant can *prove* this thesis. But Kant is quite explicit that the thesis of radical evil has *not yet been proven* (R 6:25). Presumably that is the reason that Kant concludes that the thesis is of no use in moral *dogmatics*, though he proposes that it can be used in moral *discipline* (R 6:50–51). In other words, although not theoretically established, it is plausible enough to be assumed for the purpose of our strivings to become better human beings. Those who seek for Kant’s “proof” of the thesis of radical evil, or who attempt to provide the “missing” proof, are therefore chasing an ignis fatuus. We will return to this issue in §10 of the present chapter, when we discuss Kant’s use of the thesis. But first we must try to understand the thesis itself.

3.1 ORIGINAL SIN AND RADICAL EVIL

The best place to start in discussing the thesis of radical evil is to see it as the first “fragment” of alleged (Christian) revelation that Kant chooses for his second experiment. This fragment is the Christian doctrine of Original Sin. It is a natural fragment with which to begin because Christian apologetics typically begins with our human sinfulness, from which Christianity offers us the hope of salvation. Kant’s aim in part one of the *Religion* is to offer an interpretation of the revealed doctrine of Original Sin that accords with rational morality. This can then be represented as the way a fragment of the outer circle of revealed doctrine leads back to the inner circle of pure rational religion.

“Sin,” as Christianity understands it, denotes a separation of the human being from God, especially that separation caused by the human being’s disobedience to God. Accordingly, by “sin” Kant understands “the transgression of the moral law understood as a divine command” (R 6:41–42). Since religion is the recognition of all duties as divine commands (R 6:151–152), the propensity to

evil counts as a propensity to sin when viewed from a religious standpoint – including that of the religion of pure reason.

We may identify four features of the Christian doctrine of Original Sin relevant to Kant's account. Original Sin is:

1. *A single basic sin*: Original Sin is the common ground of all our particular transgressions. This was generally regarded as Adam's sin, through which sin entered the world (*Romans* 5:16–19).
2. *Universal*: Original Sin afflicts every human being without exception.
3. *Innate*: According to Christian doctrine, Original Sin is inherited by each of us from Adam.
4. *Ineradicable*: No act of ours can remove Original Sin. This became Christian orthodoxy in 415, when the Council of Diospolis agreed with St. Augustine (354–430) in rejecting the heresy of Pelagius (354–420).

The doctrine of the radical evil in human nature, which belongs to the religion of pure reason, does not preserve these features of the revealed doctrine in their original form. Specifically, it omits the claim that sin is literally inherited by human beings from their ancestors (R 6:40). It rejects this idea on the ground that in order to be imputed to us, all evil must be our own deed (R 6:21 n, 25, 29, 32, 37, 43, 44). It proposes to locate the origin of sin not in time but only in reason (R 6:39–44). It is crucial to appreciate, however, that Kant does not *reject* the revealed account; it still belongs to the outer circle of revealed religion. But Kant treats the scriptural account of Adam's fall as a “story,” a symbolic representation calling for rational interpretation (R 6:41–43, cf. CB 8:110, R 6:79). The Kantian thesis of radical evil preserves a version of all four features of the Christian doctrine of Original Sin as part of rational religion:

1. *A single basic (radical) evil*. For Kant “radical” refers to the common *root* (Latin *radix*) of all evil acts – of whatever degree, ranging from the slightest moral transgression to the most egregious.¹ It locates their common ground in a “radical propensity to evil” or “radical perversity in the human heart” (R 6:29, 31, 32, 37, 38).
2. *Universal*: The radical propensity to evil is considered *universal* or woven into human *nature* (R 6:21, 22, 25, 29).
3. *Innate*: When Kant calls the radical propensity to evil “innate,” he does not mean that it is present in us at or prior to birth or especially that birth is in any sense the cause of it (R 6:22, 25, 32, 38). Because the propensity is

¹ Like the Christian doctrine of original sin, Kant's thesis of radical evil directs attention primarily to evil in ourselves and at least as much to minor everyday sins as to extreme forms of evil perpetrated by notorious wrongdoers. Today, however, philosophers often use the term “radical evil” to mean extreme, horrific, or egregious evil. Drawing on Kant, they then in effect attack him for discussing a different topic from theirs and usually take what he says as a failed attempt to answer questions he never asks. They seem not to care about the misunderstandings presupposed by what they are doing. Two examples: Bernstein (2002) and Card (2002).

imputable, it must have been *acquired* by the individual's free agency (R 6:21–22, 25). Thus Kant denies that *nature* – i.e., our natural inclinations – is responsible for it (R 6:21, 34–35, 57–58, 93). It is *innate* only in the sense that there is no time in a human being's life when the propensity to evil is not already present (R 6:25, 38, 42–43).

4. *Ineradicable*: That the propensity to evil is *inextirpable* (*unvertilgbar*) (R 6:51; cf. R 6:37) means that it cannot be done away with by us once and for all through any act or acts we may perform (R 6:31, 67, 72). But it can and should be *combatted* by us – perhaps only with divine assistance, but it is unknowable by us whether God's help is necessary (R 6:44–45).

Kant's thesis of radical evil is by no means merely the proposition that we have incentives of inclination that might conflict with the demands of morality. That might be true of us even if we possessed a holy will. Further, an imaginable species of rational beings could fail to be holy – it could be susceptible to committing moral evil – without having any positive *propensity* to evil, still less a propensity that is universal to the nature of the species, also innate and inextirpable.

3.2 PRACTICAL FREEDOM

Evil action is a kind of free action. Therefore, to understand Kant's treatment of it we need to understand Kant's approach to the concept of freedom. This topic is too controversial to settle finally here. So I now briefly present and defend the account I will assume in this book.

Kant's consistent position from at least the *Critique of Pure Reason* onward is that the capacity for action under moral principles, and even the capacity of logical judgment, presuppose what Kant calls practical freedom (G 4:448; cf. KrV A547–558/B575–586, KpV 5:28–29, MS 6:221). Practical freedom is the capacity to determine your own judgments and actions through reason, rather than having them necessitated by empirical causes, such as sensible impulses. Kant argues that even to raise a theoretical question, such as the question whether you are free, you have to presuppose you are free because you presuppose that you might answer it in any one of a plurality of different ways, all of which are really possible for you, and it is up to you alone which one you choose. Moreover, what you presuppose about yourself you presuppose about every other rational being as well when you observe them as able to act according to reason.

Kant thinks practical freedom in human beings can be proven empirically, through the demonstrated human capacity to judge, choose, and act independently of necessitation by sensible impulses (KrV A802/B830). He holds that this capacity has certain empirical conditions and limitations; for example, it is either nonexistent or limited during immaturity or when someone is insane or otherwise rationally incapacitated. As a matter of psychology, it can be determined *empirically* when such conditions do and don't obtain (Anth 7:213, VM 27:559).

Kant argues, however, that there remains a metaphysical problem concerning whether what appears to us to be freedom may at a deeper metaphysical level really be natural necessity, so that the practical freedom we must ascribe ourselves might be really an illusion (KrV A803/B831). This is because *practical* freedom presupposes *transcendental* freedom: the causal power to begin a state, or a series of states, entirely from oneself (*von selbst*), therefore independently of the series of natural causes (KrV A446/B474, A534/B562, KpV 5:3, 29). No *theoretical* conception of any transcendently free cause is possible for us. In response to this metaphysical problem, the most Kant thinks we can ever know is that freedom and natural necessity do not logically contradict each other (KrV A557-558/B585-586). Hence even the real possibility of freedom can neither be proven nor disproven theoretically; we can and must accept the real possibility *and the actuality* of our freedom but only for practical purposes.

Kant's claims about freedom are sometimes expressed in terms of our membership in two different "worlds." What this means is only that we can accept it without logical contradiction because we can consider our actions under two distinct concepts: an empirical or phenomenal one, in relation to which they are subject to the necessary mechanism of nature, and an intelligible or noumenal one, insofar as they are subject to the normative laws of reason and especially of morality. This enables us to consider our actions from two different standpoints, which do not logically conflict (G 4:450-454). That is the full extent of our membership in "two worlds." We do not need to believe (or Believe) in a noumenal world outside space and time.

The notion that Kant is committed to some positive metaphysical theory of freedom as present in a transcendent, noumenal realm beyond space and time is one of the most common and persistent misunderstandings of his philosophy.² Kant toys with such mere logical possibilities as part of his argument that freedom and natural causality are logically consistent. But he does not and could not consistently offer this theory as anything we can know or even Believe.

² For one prominent exposition of this account, see Watkins (2005), chap. 5. See also Adams (2018), pp. ix-x; and Allison (1990), p. 26. But from the practical presupposition of freedom, there could never follow any claims of transcendent metaphysics about how we are free. Even I have been understood as trying to provide such a metaphysical account, based on Wood, 1984, pp. 73-101 (e.g., the account of my views given in Hudson, 1994). Such misrepresentations have totally ignored the following passage from my essay "Kant's Compatibilism": "Kant is not committed to his theory of the case as an account of the way our free agency actually works. Indeed, Kant maintains that no such positive account can ever be given. Kant does not pretend to know how our free agency is possible, but claims only that the impossibility of freedom is forever indemonstrable" (Wood, 1984, p. 99). If the title of Allison's (1990) *Kant's Theory of Freedom* were meant to refer to a supernatural metaphysical theory required for Kant's moral philosophy, then that title would be nonreferring. That has always been my position. It still is.

Especially contrary to Kant's position is the thought that we are causally determined in the world of appearance but free only in some other (transcendent noumenal) world. For then the phenomenal self would be sensitive but not free, and the noumenal self would be free but not sensitive.³ That would make it impossible for either self to be *arbitrium sensitivum sed liberum* (sensitive but free), as Kant says our will is (KrV A534/B562). If Kant were committed to such a metaphysical theory of freedom, then critics would be entirely correct to observe that so-called freedom in a supernatural (even nonspatiotemporal) world could not be anything like what we take practical freedom to be. A freedom that could be exercised only in an incognizable noumenal realm outside space and time is not only something Kant could not assert consistently with critical principles. Still further: such free agency, if it pleases us to call it that, could also have no natural or social context; it would make no sense to see it as limited or conditioned empirically in the ways Kant takes it for granted our practical freedom is. We could make no sense of the thought that a timeless, spaceless, and supernatural free agency might be attracted by empirical incentives, which we have only as natural sensible beings.

In short, Kant's position on freedom is, and has to be, one of utmost metaphysical austerity. We cannot prove that freedom is impossible but also cannot explain, comprehend, conceptualize, or gain any kind of metaphysical insight into *how* it is possible (G 4:459). Kant could not self-consistently even make sense of the claim that we are free "as things in themselves." His position about things in themselves is that although we can think them, we cannot cognize them (KrV B xxvi); a noumenon is an *ens rationis*, "a concept without an object," one of the four kinds of "nothing" (KrV A290–292/B347–348). About transcendental freedom as causality "from itself" (*von selbst*) Kant denies that it can be comprehended (*niemals begriffen*) (G 4:459) – in other words, we can form no concept of its *real* possibility. A fortiori, no transcendent theory of freedom could be such that its acceptance is practically necessary. If Kant did put forward such a theory, he could not claim it to be more than logically possible and therefore could not, consistently with the critical philosophy, justify positive assent to it even on practical grounds.

³ Such errors led to the charge, first made by Reinhold, later by Sidgwick, that Kant can consider us free only when we follow the moral law, not when we disobey it. It even led Schleiermacher, in his 1799 review of Kant's *Anthropology*, to claim that Kant cannot consistently engage in the inquiry he calls "pragmatic anthropology" because Kant's anthropology proposes to investigate empirically our free agency in the sensible world (Schleiermacher, 1998). Among the many criticisms of Kant when he is taken to be committed to such a theory, see Williams, 1985, pp. 64–65; and Blackburn, 1998, pp. 246–248. If Kant were committed to a supernaturalist theory involving timeless noumenal choice outside the empirical world, then these critics would clearly be right. The Kantian position on freedom is consistent with the strictures of the critical philosophy only if it can be disentangled from such transcendent speculations. Kant's actual position regarding the traditional problem of free will leaves us with the perplexity that freedom is incomprehensible, and therefore forbids any solution to the problem compatibilists usually attempt to solve; but it also avoids any absurd commitment to a transcendent noumenal world.

In its modern form, the earliest consideration of the compatibility of freedom with a possibly alien causality did not have to do with the relation between freedom and natural causality but rather with the relation of God's grace or predestination to human freedom. As we will see in Chapter 6, Kant's position about divine causality and human freedom is the same total agnosticism as his position about natural causality and freedom (R 6:191; cf. R 6:143–144, KpV 5:100–101, VpR 28:1110). For Kant, in short, in its moral and religious significance, human freedom is never a matter of metaphysics or supernatural causality in a noumenal “beyond.” The truth about our maxims and disposition that is hidden from us is not a transcendent metaphysical truth but an empirical truth about ourselves that is concealed through our self-deception and revealed (as far as it can be) only through our efforts at self-honesty.⁴

3.3 THE CONCEPT OF MORAL EVIL

In part one of the *Religion*, we are not talking about “evil” in the sense of the bad things that happen to people. Kant would call that “ill” (*Übel*). We are not even talking about “systemic” social evils for which no individual bears responsibility, such as “general injustice” (VE Vigilantius 27:416, Ak 19:145; see Wood, 2008, pp. 198–200, and Wood, 2014, pp. 83–89).

Evil (*Böse*) is a deed (an action falling under obligation) done freely and culpably by a morally responsible individual agent. An action, as Kant understands the concept, is something within the power of a rational agent that the agent freely chooses as a means to an end to be produced, an end the agent also freely sets for itself. Actions are therefore always done for reasons: an action is always done for the sake of an end; it follows a maxim (or subjective principle) that the agent has freely chosen. If evil were not an action, done for reasons, it could not be imputable to human beings, and therefore it could not be evil. But we cannot suppose that the reasons for which evil is done are rationally sufficient because the moral reason an agent has not to do evil necessarily overrides any reason the agent might have to do it. If there weren't rationally decisive reasons not to do evil, then there could be no ground to say that it is something we could be morally required *not* to do. Blaming the agent for it would be irrational.

By its very concept, an action can be made intelligible only in terms of the reasons for which it is done. Yet an evil action is one there are decisive reasons not to do. This means that the concept of evil is paradoxical, and an evil act cannot in principle be made wholly intelligible. Kant is well aware of this point; it is a theme emphasized repeatedly (almost obsessively) in the *Religion* (R 6:21, 25, 32, 35, 41, 43).⁵ Evil, he says, is “inscrutable” (*unerforschlich*) (R 6:21; it

⁴ A recent study that sheds much-needed light on this theme in Kant is Papish (2018).

⁵ Not appreciating this point, or drawing the wrong conclusion from it, some commentators complain that Kant's thesis of radical evil “fails to explain anything.” Others speak of Kant “confessing defeat” in the attempt to “explain evil.” See William McBride, “Liquidating the Just

must be attributed to “the free power of choice, . . . but there cannot be any further cognition of [its] subjective ground or cause (although we cannot avoid asking about it)” (R 6:25). Evil is a form of motivated unreason, like *akrasia*, *acedia*, or self-deception; all three of these Kant considers to be forms of evil or at least closely related to it.⁶ Kant does think, however, that we can form a coherent concept of evil; we can even understand it, at least up to a point, both as a human propensity and in relation to the development of the human species in history.

Two bad ways to look at evil. There are two common ways, both explicitly considered and rejected by Kant, of trying to make evil *seem* more intelligible than it ever can actually be. The *first* way is to locate the ground of evil in our natural inclinations themselves. An explanation of evil based on natural causality would remove evil from the realm of free choice and moral imputability, which would do away with it *as evil*. The alternative version of this way, the thought that animal or sensuous nature is morally culpable or a matter of reproach, is frequently attributed to Kant himself, based on passages where he cites inclinations as obstacles or a counterweight to the moral law (G 4:405, KpV 5:21–28, MS 6:376, 379–380). In the *Religion*, however, Kant’s rejection of this confused thought is explicit, repeated, and emphatic: “The ground of evil cannot (1) be placed, as is commonly done, in the sensuous nature of the human being, and in the natural inclinations originating from it” (R 6:34–35, cf. 6:21, 26–27, 31, 36, 57–60). Evil, Kant says, cannot lie in natural inclinations that “openly display themselves unconcealed to everyone’s consciousness, but is rather as it were an invisible enemy, one who hides behind reason and is hence all the more dangerous” (R 6:57). Kant’s position is that inclinations can be involved in evil only when they have been incorporated as incentives into a freely chosen evil maxim. Then it is this *maxim* that is evil, not the inclinations themselves.

The *second* way to give evil a false appearance of intelligibility is to represent it as something that a human being might have a genuine and sufficient reason for doing. One variant of this strategy, not dealt with directly by Kant, involves restricting the scope of reason to the theoretical or to instrumental or prudential reason, and then thinking of an action as “evil” can be only our negative attitude to actions that might be perfectly rational by these standards. This view runs aground on the objection that it could give us no way of rationally dissuading others from evil action and therefore no rational justification for

Society: Radical Evil’s Triumphant Return,” in Schrift (ed.), 2005, p. 38; and Bernstein (2002), p. 33. Compare the better treatment of these issues by Robert Louden: “Evil Everywhere: The Ordinarity of Radical Evil,” in Anderson-Gold and Muchnik (eds.), (2010), p. 96.

⁶ Papish (2018) is right to explore self-deception as lying at the root of evil for Kant. Her book is valuable for exploring this theme. I am puzzled, however, by many things in it, including her claim (p. 4) that I would be among those who are *skeptical* about the claim that for Kant evil involves self-deception. I don’t know where she got that impression, since it is quite false.

condemning it as evil. Besides, Kant belongs to a long tradition that regards morality as reason's highest function. Evil action may meet instrumental or prudential standards of rationality, but it is evil precisely because it is contrary to the highest (the moral) standard. To deny that this is the highest standard is in effect to say that there is no such thing as evil – only some things we happen to dislike or of which we happen to disapprove.

Another variant of the second strategy, which Kant takes up explicitly, is to ascribe to the human will itself an original rational incentive to disobey the moral law. Kant's rejection of this alternative, of a "diabolical will" or an "evil reason," is equally emphatic: "The ground of evil can also not be placed (2) in a corruption of the morally legislative reason." For, Kant argues, it would be contradictory to think of a rational being as bound by the moral law yet either lacking any rational incentive to obey it or possessing a rational incentive to disobey it. Related to this is Kant's denial that the corruption of our will results in the total loss of our moral predisposition. For that would make our acts no longer imputable, hence, not evil; and if we once lost the capacity for good, we could then never do anything to recover it (R 6:46–52). "*Sensuous nature* contains too little to provide a ground of moral evil in the human being," since it constitutes the human being only an animal and not a moral being. Yet "a reason exonerated from the moral law, an *evil reason* (as it were) . . . contains too much, because resistance to the law would thereby be elevated to an incentive," and this would again contradict the imputability of evil (R 6:35).

Kant's rejection of the "diabolical will" is sometimes criticized on the ground that it excludes the possibility (as the critics like to put it) of "doing evil for evil's sake." Since we know this is something of which human beings are deplorably capable, critics then represent Kant as sadly oblivious to the full depths of human evil.⁷ But this criticism is grounded on a basic misunderstanding. Some philosophers hold that all incentives must be contingent and empirical so that incentives to do what is right and incentives to do what is wrong are on a par. But Kant holds that reason provides us with an original rational incentive, grounded in our predisposition to personality, to obey the moral law (R 6:28). It would be self-contradictory, however, for a *rational* faculty to have an original incentive to choose the opposite of what it is unconditionally rational to choose. When he denies the possibility of a diabolical will, that self-contradiction is what Kant is rejecting as impossible. Moreover, no will could reasonably be considered *evil* (blamable) for doing precisely what it had an original rational incentive to do. Such a will would be, as Kant says, "exonerated from the moral law" (R 6:35).⁸

⁷ This is the criticism of Kant presented long ago by John Silber (1960) and reiterated more recently: Silber (1991), pp. 177–211; Silber (2012), pp. 329–334. See also Card (2002), p. 212; and Matušík (2008).

⁸ Kant apparently does subscribe to the scholastic adage *Nihil appetimus, nisi sub ratione boni; nihil aversamus, nisi sub ratione mali* ("We desire nothing except under the reason of the good; we avoid nothing except under the reason of the bad"). His reservations about this adage, expressed at KpV 5:59, have to do with an ambiguity he finds in the words "good" and "bad," not with the

Perhaps in demonizing those they hate, some people do frame the incoherent notion of a will that is (as they might put it) “*so evil*” that it is incapable of good and constituted solely by the incentive to do evil. However, this could only be an incoherent projection of their irrational malice toward those they hate, whom they are now depicting to themselves in nonsensically venomous terms.⁹

Kant certainly holds that there are “diabolical *vices*” – the “vices of hatred”: envy, ingratitude, and malice – that set the unhappiness of others as an end, even though it is our duty to include the happiness, never the unhappiness, of another among our ends (R 6:27, MS 6:385–388, 458–461). These vices do involve the knowing choice of an evil end and the choice of it *as* evil. Kant allows, in other words, that people may have an *inclination* to pursue an evil end, *qua* evil, simply for its own sake. He sometimes describes these vices as “exceeding human nature” (R 6:35–36, MS 6:461). But he is not denying that such vices occur; he is rather claiming that they are exceptional, falling outside the normal range of human nature.¹⁰

Kant holds that in order to do evil, we must represent to ourselves something evil as good, perhaps even evil itself as good. He would infer that all evil is grounded on self-deception, which makes the lie to oneself for Kant the most fundamental form of evil (MS 6:431, R 6:42 n). Kant admits that radical self-deception may be hard to conceptualize because it is inscrutable in precisely the way evil itself is; but it is unquestionably exhibited by human beings.

Kant’s basic conception of evil. Evil is a matter of the two basic kinds of rational incentives present to human action: (1) the moral incentive, to conform our choices to duty or the moral law; and (2) nonmoral incentives, which Kant characterizes in general terms as those of inclination or self-love. There is a clear rational priority: the moral incentive prevails decisively over all others. Evil choice can be expressed as a *maxim* (or generalized intention) of subordinating the moral incentive to some rationally weaker incentive of inclination or self-love.

The difference, whether the human being is good or evil, must lie not in the difference between the incentives that he incorporates into his maxim (not in the material of the maxim), but in their subordination (in the form of the maxim): which of the two he makes the condition of the other. It follows that the human being (even the best) is evil only because he reverses the moral order of his incentives in incorporating them into his maxims (R 6:36).

substantive thesis asserted in the adage itself. The “guise of the good” thesis (as it has recently been called) was first challenged in the fourteenth century by William of Ockham. It has recently been controverted by David Velleman (1992) and Kieran Setiya (2007). The thesis has been defended in qualified form by Raz (2010) and by philosophers with Kantian leanings: Tenenbaum (2007) and Sussman (2009).

⁹ This point is made effectively by Caswell (2007), pp. 147–157.

¹⁰ Kant’s diabolical vices might be “radical evil” in the now popular sense of “egregious evil.”

Kant does not aim at reducing evil volition to some common stereotype – such as the self-indulgent hedonist or the self-interested schemer.¹¹ Since his purpose is to provide the most abstract and encompassing concept of evil, “inclination” and “self-love” should be understood as merely placeholders for whatever nonmoral incentives might be chosen in preference to those of morality. “Inclination” refers to any habitual empirical desire, whatever its object might be. “Self-love” refers to the rational desire of prudence associated with our predisposition to humanity, to promote whatever we consider to belong to our happiness or well-being, whatever content we may include under this heading.

Kant’s account tends to disappoint those who want from a philosophical account of evil the deliciously lurid evocation of some uncanny state of mind that it pleases us to imagine motivating the most monstrous evils we can think of. Wiser observers have pointed out that the greatest evils, such as the Holocaust, may sooner result from “total banality and prosaic triviality” than from some “satanic greatness” (Arendt, 1994, p. 459; cf. Arendt and Jaspers, 1992, especially letters 337–345). Even the greatest evil is small, pitiful, and contemptible.

3.4 THE PROPENSITY TO EVIL

Kant speaks of the radical, universal, innate, and inextirpable evil in human nature as a “propensity” (*Hang*). He distinguishes a *propensity* from a “predisposition” (*Anlage*). A predisposition is a capacity (natural or acquired) involving a characteristic tendency to certain kinds of behavior, typically behavior with certain specifiable results, usually beneficial to an organism. Some predispositions are placed in us by nature. Many, including talents and skills, are developed by us through our own efforts (I 8:18–20). Natural predispositions cannot be imputed to us as blamable, and Kant insists that natural predispositions are without exception good in the sense that they are placed in us by nature *for* good. Kant distinguishes three basic natural predispositions that are the ground of all other natural human predispositions:

1. *Animality*. This includes the prerational or instinctive tendencies in our nature to self-preservation, preservation of the species (sexuality), and sociability (a prerational tendency to associate with others of our kind). This includes purely animal self-love (*Selbstliebe, solipsismus*) (KpV 5:73).

¹¹ Some charge Kant is doing such a thing (see Card, 2010, p. 76; and Bernstein, 2002, p. 42). This is belied by some of Kant’s most conspicuous assertions about the worst forms of evil. These are the “diabolical” vices, or vices of hatred: envy, ingratitude, malice, and vengeance are the worst that human nature has to offer (R 6:458–461, R 6:27). As evidence for the evil in human nature, he prefers to cite not self-interested behavior but “never-ending cruelty . . . from which no human being derives the least benefit” (R 6:33). Or lying and deception (especially betrayal of another’s trust) (MS 6:429–431, 437–440, R 6:33–34). All these graver forms of evil are most often displayed by people who are not rationally self-interested but even conspicuously irrational even by nonmoral standards: deluded and imprudent.

It is in itself good, but through the exercise of our freedom, certain vices (those of brutishness and savagery) can be “grafted onto it” (R 6:26–627).

2. *Humanity*. This term includes two predispositions: the *technical* predisposition – the rational, but nonmoral, tendency to set ends according to reason, to choose actions as means to them (Anth 7:322–323) – and also the *pragmatic* predisposition to unite the ends of inclination into a single encompassing end, under the name of one’s own happiness (Anth 7:323–324). In the *Religion* Kant emphasizes that both forms of the predisposition to humanity make use of the rational capacity of *comparison*. We might think that he means only the capacities to compare different means to our ends and also to compare ends regarding their contribution to our general well-being. Clearly both would be essential to the predisposition to humanity. But Kant emphasizes that humanity also involves a quite different kind of comparison: of myself with others. So regarded, natural self-love (*Selbstliebe*) assumes a different, specifically rational form: *Eigenliebe*, *philautia*, which is associated with an evaluation of oneself and a propensity to place a higher value on oneself than one does on others, which is *self-conceit* (*Eigendünkel*, *arrogantia*) (KpV 5:73). This propensity is not naturally implanted but acquired through the free exercise of our rational faculty. Kant says that the “vices of culture, envy, ingratitude, joy in another’s suffering (*Schadenfreude*)” and the specifically “diabolical vices” can be “grafted” onto it. We will see presently that this comparative-competitive form of rational self-love is closely associated in Kant’s account with the radical propensity to evil (R 6:27).
3. *Personality*. This is the moral predisposition, the capacity to make the moral law a sufficient incentive to action. It includes moral feeling and grounds the imputability of actions. Nothing evil can be grafted onto this predisposition (R 6:27–28). Evil consists in preferring some other incentive over the moral incentive, which is offered by personality.

In contrast to a natural *predisposition*, a *propensity* is “the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*) insofar as this possibility is contingent for humanity in general” (R 6:29).¹² It might be supposed that the two terms put in parentheses here are being equated with each other, perhaps both also with “inclination.” But closer consideration shows that this is not so: “habitual (empirical) desire” is the same as “inclination,” but *concupiscentia* refers back to the phrase “possibility of an inclination.” A brief explanation of these matters will prove clarifying.

¹² The concept *concupiscentia* is part of the traditional doctrine of Original Sin, as developed by St. Augustine and others. Its scriptural source is the attribution to human beings of *yetzer hara* (יֵצֶר הָרָע): “And God saw that the wickedness of man was great in the earth, and that every imagination of the thoughts of his heart was only evil continually” (Gn 6:5, King James version, cf. Gn 8:21; Kant uses the Lutheran translation of this phrase – *Tichten und Trachten* – at G 4:407).

An *inclination* (*Neigung*) for Kant is an empirical desire but not just any empirical desire: it is a *habitual* desire – that is, a desire involving a maxim on the part of the agent, a desire to seek a certain *kind* of satisfaction through a certain *kind* of action (Anth 7:251). Desire itself, according to Kant, at least in human beings, is itself a “self-determination of the subject’s to bring about through a representation the object of that representation” (Anth 7:251). For beings like ourselves, desiring itself is already an act of freedom. In other words, our will is *responsible* for our desires; for Kant as for Aristotle, virtues and vices can consist in part in our inclinations and in what pleases or pains us (Aristotle, *Nicomachean Ethics* 1107a–b). For Kant an envious person is one who desires the misfortune of rivals and has a propensity to view their well-being with distress (MS 6:458). If I am an arrogant person, I desire that others think less of themselves when they compare themselves with me (MS 6:465). But inclinations are not imputable to us as *deeds* (i.e., as possibly deserving reward or punishment) because actions are not deeds unless they involve the actual setting of ends and the choice of means and also fall under obligatory laws (MS 6:223). A *propensity*, however, is not yet even an action, still less a deed; it is rather a “subjective determining ground of the power of choice that *precedes every deed* and hence is not yet itself a *deed*” (R 6:31).

A propensity is a “concupiscence” (*concupiscentia*, *Concupiscenz*): “Concupiscence (a lusting after something) must be distinguished from desire itself, as a stimulus [*Anreiz*] to determining desire” (MS 6:213, cf. MS 6:393).¹³ Kant also identifies one other important feature of a propensity: it is “the predisposition to desire an enjoyment which, when the subject has experienced it, arouses an inclination to it” – for example, the predisposition once one has drunk alcohol to desire to drink more (R 6:29 n). In this respect, we may compare (and contrast) a *propensity* with an *instinct*. For an instinct is a natural predisposition to desire or be pleased by something even prior to acquaintance with it (Anth 7:265; VA-*Menschenkunde* 25: 1111–1112, VA-Mrongovius 25:1339, VA-Busolt 25:1518). We desire something instinctively when we desire it as soon as we become acquainted with it, based on a preexisting natural predisposition to desire it. But a propensity is a tendency to desire that is acquired only after acquaintance with the object desired. When Kant describes the propensity to evil as “contingent for humanity in general,” he clearly intends this to mean that it is not something implanted in us by nature but something arising in us from the exercise of freedom.

¹³ This seems to play a role in Kant’s conception in part two of the *Religion* of the “personified idea of the good principle” (the “Son of God,” “the ideal of humanity well-pleasing to God,” or “the prototype of holiness”) – in other words, the Christ image (R 6:60–66). See Chapter 5 below. Kant wants to preserve the full humanity of Christ, including his being subject to the same temptations as we are, while nevertheless representing this ideal as a holy will: “unchanging purity of will, not gained through effort but innate, [rendering] any transgression on his part impossible” (R 6:64, cf. KpV 5:32). So this human ideal of holiness is apparently afflicted innately with the *propensity* to evil but is not susceptible to actually *committing* evil.

3.5 KANT'S "RIGORISM"

It seems clear that one reason Kant seems disposed to accept the thesis of radical evil is his methodological commitment to beginning with a fragment of alleged revelation. In addition, however, Kant offers rational grounds for thinking that the thesis of radical evil might be plausible: it is made more so by the position he describes as "rigorism": "To preclude as far as possible anything morally intermediate, either in actions (*adiaphora*) or in human characters" (R 6:22). But before we consider its use in his arguments, we must be cautious in interpreting the position Kant is endorsing under this name and appreciate his arguments in favor of that position.

The qualification "as far as possible" is significant. Rigorism is certainly *not* the claim that there is nothing morally indifferent or morally ambiguous. Kant regularly insists there are moral *adiaphora* both in human actions (MS 6:223, 458) and (right in the *Religion* itself) in matters of religious faith and practice (R 6:44 n). He directly condemns the refusal to allow anything morally indifferent, describing this as a conception of "fantastic virtue" and saying that such a view would transform the dominion of virtue into a tyranny (MS 6:409). Kant is an outspoken foe of any attitude of "moral clarity" that divides people into "good" and "bad." He is suspicious of our tendency to see some people as "diabolical" and others as "angelic" (MS 6:461). When we consider others whose actions we condemn, Kant would encourage us to see how much we may have in common with them and how our guiltlessness may be merely good fortune in having escaped their temptations (MS 6:392–393).¹⁴

Further, Kant's rigorism in the *Religion* is narrow in scope, applying only to maxims that incorporate the moral incentive along with other competing incentives. Kant's rigorism is then supported by a simple and cogent argument, proceeding from the thesis that "the moral law is itself an incentive in the judgment of reason" to the conclusions that, *first*, "whoever makes it his maxim is morally good" and, *second*, whenever an agent incorporates some other incentive into his maxim *ahead of* the moral law, this can be considered neither morally good nor indifferent but must be judged evil (R 6:24). That is the full extent of the *Religion*'s "rigorism."¹⁵

Clearly Kant recognizes many actions (the ones he counts as *adiaphora*) where no such choice is required. Indeed, in the case of some dutiful actions

¹⁴ Card (2011) pp. 77–92, provides a long list of "moral intermediates," as though they were counterexamples to Kant's rigorism. The list itself, and her discussion of it, displays a lot of moral and philosophical insight. But the use of it to criticize Kant is misplaced because Kant's "rigorism" allows for many moral indifferents, intermediates, and ambiguities.

¹⁵ Moral philosophers influenced by Kant often use the term "rigorism" in a very different sense, one never used by Kant himself. They use it, namely, to refer to the doctrine (commonly but mistakenly ascribed to Kant) that moral rules admit of few or no exceptions. Then the issue for a "Kantian" is whether and how to make exceptions to rules. See, for example, Schapiro (2006). But Kant holds no position at all that is rigorist in this sense. See Wood (2017), pp. 46–48.

no such choice is ever in question – namely, the kinds of action Kant considers early in the *Groundwork*, where duty and some other nonmoral motive (prudence or immediate inclination) both provide sufficient incentives to act. Since they conform to duty, such actions are good, deserving praise and encouragement. They are in no way evil or even indifferent, precisely because they offered no choice between the incentive of morality and incentives of self-love or inclination (G 4:397–399). In part one of the *Religion*, however, the question is whether a position of indifference is possible for a human being or for a human character. His answer seems to be that each of us, at the deepest level, is constantly confronted with the question which incentive is to be preferred or made the condition of the other; about this we can in principle never be morally indifferent (R 6:36). In regard to this choice between moral and nonmoral incentives, our “disposition” or “fundamental maxim” is always either good or evil, never indifferent.

Kant thinks any such judgment is never unproblematic, however, due to human self-opacity and self-deception. We can never judge it directly through the empirical observation of actions.

We cannot observe maxims, we cannot do so unproblematically even within ourselves; hence the judgment that an agent is an evil human being cannot reliably be based on experience. In order, then, to call a human being evil, it must be possible to infer *a priori* from a number of consciously evil actions, or even from a single one, an underlying evil maxim, and from this, the presence in the subject of a common ground, itself a maxim, of all particular morally evil maxims (R 6:20).

Scholars have been tempted to read this passage as alluding to some sort of “a priori” or “transcendental” argument for Kant’s thesis of radical evil in human nature. Read carefully, however, it suggests nothing of the kind, and we see that such a reading completely misunderstands its point. In the opening paragraphs of part one, Kant is considering two opposing views: the biblical view that “the world lieth in evil” (John 5:19) and the more “heroic” modern view that human beings have a predisposition to good (R 6:20). In part one of the *Religion*, we find that between these two views, Kant’s sympathies are divided. In his attempt to devise a way of clarifying the issue, he is here asking, first, simply *what it means to judge that a single human agent is evil*. There is as yet no question about Kant’s thesis that there is radical evil in human nature. No issue of universality, innateness, or inextirpability has yet even been hinted at in this passage.

Further, Kant’s claim that the judgment that a human being is evil “cannot be based on experience” obviously cannot mean that the judgment is wholly a priori in the way that a transcendental argument would have to be. Kant’s point is rather that we are not entirely reliable judges of our own maxims or intentions, especially of the deeper and more pervasive ones. We therefore cannot decide whether someone’s fundamental maxim is good or evil merely on the basis of empirical reports of their good and evil actions. The inference he

countenances is said to proceed from one or more “consciously evil actions” – whose occurrence could plainly be known only by experience, never a priori. Kant therefore has to be speaking here only about the relative kind of apriority that attaches (for example) to the judgment that if the foundation of a house is undermined, then it will collapse (KrV B2). His point in the passage is that any judgment that someone is a good or evil human being must rest on an a priori connection between the person's empirical actions and an underlying maxim presupposed by them.

3.6 WALPOLE'S THESIS

“Every man has his price.” This is a thesis often attributed to Sir Robert Walpole (1676–1745), who is generally regarded as the first prime minister of England. Kant quotes it in part one of the *Religion* (R 6:38–39). When Walpole said this, he may have been referring not to all men but only to certain of his political opponents (see Coxe, 1816, 4:369). But when Kant quotes this saying in part one, he intends it to apply to all human beings. What it means is that for any moral agent, and any duty that agent may have, the agent would violate that duty if either the price offered him for that violation in terms of his own self-interest were high enough or the cost to him of doing his duty were sufficiently great.

Walpole's thesis, as I will call it, might be looked upon as a corollary, or perhaps an extension, of the claim that the radical propensity to evil is *inextirpable*: we can't get rid of it once and for all. Walpole's thesis says that no human being will ever attain so high a degree of moral virtue that they would not still yield to some temptation if only the temptation were great enough. This perhaps extends the inextirpability of the propensity to evil because it says not merely that we will always *have a propensity* to choose nonmoral incentives over the moral incentive but that we will always *actually make that choice* if the nonmoral incentives (whether positive or negative) are sufficiently great. Of course as long as we remain free agents, to whom deeds may be imputed (i.e., as long as we have what Kant calls “freedom in the negative sense”), we will still always *be able* to resist the nonmoral incentives and do what we ought. This is the meaning of the Kantian thesis usually expressed as “Ought implies Can.”¹⁶ What Walpole's thesis says is that if the price of duty is sufficiently high for us, we *will not choose* to do what we nevertheless still *can* do.

Kant never quite endorses Walpole's thesis. He says he leaves it to each reader to decide whether to accept it; but he draws from it the following Pauline claim, which he does seem to accept, at least as part of the thesis of radical evil: “There is no distinction here, they are all under sin – there is none righteous (in the spirit of

¹⁶ This is the only actual meaning of the Kantian principle “Ought implies Can.” It is concerned exclusively with the relation of rational incentives to *motivational* possibility. For a thoughtful review of the Kantian texts that bear on this principle, see Stern (2004). See also Kahn (2019).

the law), no, not one” (R 6:39, cf. *Romans* 3:9–10). Kant clearly thinks that rigorous self-examination will find in every one of us consciously evil actions from which we can infer a maxim involving the (occasional) deviation from the moral law, expressing the underlying evil maxim of preferring nonmoral incentives to the moral incentive (R 6:32). That we would all find such a tendency in ourselves seems extremely plausible. But it is also quite remarkable, when we consider what it means about us as rational beings. It will help us to pause and consider this point.

In the *Critique of Practical Reason*, Kant considers a man who “asserts of his lustful inclination that when the desired object and opportunity are present, it is quite irresistible to him.” Kant then suggests that we ask him “whether, if a gallows were erected in front of the house where he finds this opportunity” on which he would “be hanged immediately after gratifying his lust, he would not then control his inclination. One need not conjecture very long what he would reply” (KpV 5:30). Kant is supposing that the rational strength of the incentive to self-preservation would so far outweigh his lascivious impulses as to make the choice obvious and easy. A man who chose to satisfy his lust under those conditions would be not only extremely irrational but also extremely rare.¹⁷ The moral law, however, is supposed to be a decisively stronger rational incentive than any incentive whatever arising from inclination or self-love. It therefore ought to strike us as extremely remarkable – a sign of something very significant about us as rational agents – that we find in ourselves a strong tendency, at least occasionally, to choose self-love or inclination over morality. We clearly cannot infer such a tendency a priori merely from the fact that we have nonmoral incentives. If that were so, we could also validly infer a priori that Kant’s lustful man has a propensity to choose lust over survival. Why, then, is the propensity to evil – to choose against a propensity rationally stronger even than love of life – so common and familiar to us that we take it for granted, just as if it were something quite rational?

3.7 EVIL AND UNSOCIABLE SOCIABILITY

Kant thinks that the result of our self-examination harmonizes with some of the views about empirical human nature that he came to adopt under the influence of Jean-Jacques Rousseau. We saw above that Kant thinks our natural predisposition to humanity involves the rational faculty of comparison, not only regarding the objects and ends of rational choice but also regarding comparisons between ourselves and others. In his *Discourse on the Origin of Inequality among Men*, Rousseau argued that human “perfectibility,” the development of human

¹⁷ Of course, the man’s irrational conduct in facing death would be less unexpected if we thought of him as moved not merely by a lustful desire for sexual gratification but by sexual *passion* (an inclination difficult for reason to control). If the man were to risk death out of *love* (which has a moral component), then his conduct might not even be contrary to reason.

faculties, including the faculty of reason, was associated with a new form of self-love. Natural or innocent self-love, which Rousseau called *amour de soi-même*, came to be accompanied, or even supplanted, by a new kind: *amour propre* (Rousseau, 1997, pp. 152–154, 161–163, 218).¹⁸ We have seen already that Kant refers to it in discussing the development of our predisposition to humanity, using two different though closely related terms: self-love (*Eigenliebe*, *philautia*) and self-conceit (*Eigendiinkel*, *arrogantia*) (KpV 5:73). Kant's philosophy of history is grounded on the thought that nature's purpose in developing human predispositions was served by the same trait, to which Kant refers using a term drawn not from Rousseau but from Montaigne: "unsociable sociability" (*ungesellige Geselligkeit*).

The means nature employs in order to bring about the development of all their predispositions is their antagonism in society, insofar as the latter is in the end the cause of their lawful order. Here I understand by "antagonism" the *unsociable sociability* of human beings,¹⁹ i.e. their propensity to enter into society, which, however, is combined with a thoroughgoing resistance that constantly threatens to break up this society. The predisposition for this obviously lies in human nature.... Now it is this resistance that awakens all the powers of the human being, brings him to overcome his propensity to indolence, and, driven by ambition, tyranny and greed, to obtain for himself a rank among his fellows, whom he cannot *stand*, but also cannot *leave alone* (I 8:20–21).

Kant makes the connection of evil with sociability explicit especially in two passages:

These vices [grafted onto our predisposition to humanity] do not really issue from nature as their root but are rather inclinations, in the face of the anxious endeavor of others to attain a hateful superiority over us, to procure it for ourselves over them . . . ; for nature itself wanted to use the idea of this competitiveness (which in itself does not exclude reciprocal love) as only an incentive to culture (R 6:27).

The human being is nevertheless in this perilous state [the radical propensity to evil] through his own fault; hence he is bound at least to apply as much force as he can muster in order to extricate himself from it. But how? That is the question. – If he searches for the causes and the circumstances that draw him into this danger and keep him there, he can easily convince himself that they do not come his way from his own raw nature, so far as he exists in isolation, but rather from the human beings to whom he stands in relation or association. It is not the instigation of nature that arouses what should properly be called the passions, which wreak such great devastation in his originally good predisposition. His needs are but limited, and his state of mind in providing for them moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and will despise him for it. Envy, addiction to power, avarice, and the malignant inclinations associated with these, assail his nature, which on its own is undemanding, as soon as he is among human beings. Nor is it

¹⁸ An excellent treatment of this theme in Rousseau is found in Neuhouser (2008).

¹⁹ "There is nothing more unsociable than man, and nothing more sociable: unsociable by his vice, sociable by his nature" ("Of Solitude," Montaigne, 1991, p. 267).

necessary to assume that these are sunk into evil, and are examples that lead him astray: it suffices that they are there, that they surround him, and that they are human beings, and they will mutually corrupt one another's moral disposition and make one another evil (R 6:93–94).

As these quotations show, to say that for Kant the radical human propensity to evil has a social and historical origin is only to report what Kant explicitly says. Some commentators, however, insist that taking Kant at his word leaves him with a doctrine that simply can't be right or at least can't be consistent with his other views. (They usually do not attribute the errors to Kant, however, but blame them on the "misinterpretation" of Kant by scholars such as myself who quote the above passages and then directly report that they say precisely what they do say.)

One objection is that to place evil in the context of human competitiveness is to exempt human individuals from responsibility for it.²⁰ A related objection is that attributing evil to our social condition is "reductionist"; it must ignore the "noumenal" context of evil, and it allegedly suggests that evil can be made comprehensible in merely naturalistic terms.²¹ Yet another common objection is that the social origin of evil can't apply to the violation of duties to oneself: so there must at least be two distinct forms of evil for Kant, only one of which is social in origin while the other belongs to the noumenal realm only and cannot involve anything social.²²

The first and second groups of scholars seem to think that for Kant the social condition must consist of nothing but external causal interactions in the world of appearance, while free action must take place somewhere else entirely ("up there in the noumenal world"). These errors have been exposed in §2 above. Kant approaches the empirical observation of human actions in their social and historical context based on the practical presupposition that we are free as members of the *empirical* world. It is as free agents in *this* world that he investigates evil action.

Regarding the objection based on self-regarding duties: to say that evil has a social context and origin is not to say that every case in which we act on the evil maxim is one in which we directly pursue superiority over others. The maxim of evil is grounded in a certain way of valuing ourselves – the comparative-competitive way. This leads us to value our *state* or *condition* (*Zustand*) over our *person*, and the happiness of the former over the virtue or morality of the latter, since (as we have seen) "we consider ourselves happy or unhappy only in comparison to others."²³

²⁰ See Grenberg, 2005, pp. 31–42; and "Social Dimensions of Kant's Conception of Radical Evil," in Anderson-Gold and Muchnik (eds.), 2011, pp. 173–194; and in the same volume, Muchnik, "An Alternative Proof of the Universal Propensity to Evil," *ibid.*, pp. 128–130.

²¹ See Morgan, 2005, pp. 110–111. ²² See Kemp, in Thorndike, 2011, pp. 100–120.

²³ Kant's account is continuous with that of Adam Smith: "It is chiefly from regard to the sentiments of mankind that we pursue riches and avoid poverty. For to what purpose is all the toil and bustle of this world? What is the end of avarice and ambition, the pursuit of wealth, of power and pre-eminence? Is it to supply the necessities of nature? The wages of the meanest laborer can

When we consider Kant's specific accounts of self-regarding duties and vices, it becomes evident that they are no exceptions to the social origin of the human propensity to evil. To relate human evil to social comparison is not to say that every evil act directly involves harming others. For civilized beings like ourselves, even "private" vices indirectly involve social comparison: for even the simplest pleasures are more delicious when we know others cannot afford to indulge in them. Other vices involving a violation of duties to ourselves more directly involve social comparison. On Kant's account, "bestial" vices (such as gluttony), and violation of our duty of self-preservation (G 4:421–422, 429, MS 6:422–424), do not directly involve social comparison, but they do involve the excessive preference for pleasure in our state over the dignity of our person (MS 6:427). Our state is valued because we can compare it to the state of others – hopefully to our own advantage (R 6:27). Lying, whether as a means of manipulating another or, in relation to ourselves, maintaining a flattering false image of who we are, is manifestly social in motivation and content. The motivation of self-regarding vices of avarice (MS 6:432–434) and servility (MS 6:434–436) involves social comparison and competition in quite obvious ways. It displays a shallow view of self-regarding vices to see them as exceptions or counterexamples to the social context and origin of evil.

3.8 THE PASSIONS

The instinctual desires arising from animality are not themselves evil, but they can have vices "grafted onto" them: the "bestial" vices of gluttony, lust, and wanton cruelty (R 6:26–27). These vices, as we have seen, manifest unsociable sociability indirectly through their preference for the worth of our state or condition (*Zustand*) over the worth of our person. A more direct manifestation is found in the vices "grafted onto" the predisposition to humanity. As we saw above (R 6:93–94), this prominently includes the development of the *passions*. Kant uses this term in a different and narrower sense from its use in some other modern philosophers, such as Descartes and Hume. Passions, in Kant's sense of the word, are inclinations that take the form of "mania" (or "addiction") (*Sucht*) because they resist comparison with and limitation by other desires and consequently resist the influence of reason (VA 7:265). Although they resist reason, the passions arise only in connection with it and presuppose a maxim (VA 7:266). Because they resist reason, the passions are also evil (VA 7:267). These passions are allied to the "vices of culture" or "diabolical" vices that may be "grafted onto" humanity: envy, ingratitude, and gloating over harm to another (*Schadenfreude*) (R 6:27).

supply them ... It is the vanity, not the ease, or the pleasure [of the higher ranks of life] that interests us ... Compared with the contempt of mankind, all other external evils are easily supported ... It is not ease or pleasure, but always honor, though frequently an honor very ill understood, that the ambitious man really pursues" (Smith, 2000, pp. 70–71, 89, 83).

All evil, as we have seen, is for Kant based on the “inner lie” (or self-deception). A striking feature of the passions is also the way in which they involve deceptiveness (self-deception) or even delusion (*Wahn*). A delusion is “the inner practical deception of taking what is subjective in a motive for something objective” (VA 7:215). A passion “presents objects of imagination as real ends (ways of acquiring honor, domination or money)”; and inclinations for these things “are apt to become passionate in the highest degree, especially when they are applied to competition between human beings” (VA 7:275). Passion afflicts the passionate person with a self-incurred inability to compare a desired object rationally with other such objects, hence also to appreciate the rational strength of the incentive that it provides in relation to other incentives (whether of inclination, prudence, or morality).

In regard to its effect on our behavior relating to other people, every passion involves a “slavish disposition, through which another, having gained power over it, acquires the capacity to use one’s own inclinations to serve his aims” (VA 7:272). (“Those who believe themselves the masters of others cease not to be even greater slaves than they” [Rousseau, 1997, bk. I, chap. 1].) The greedy are the most easily swindled; those hungry for power are the most slavish followers of the powerful; and those who most passionately seek the admiration of others tend to be the most fawning admirers of celebrities. “Here the human being becomes the dupe of his own inclinations, and in his use of such means he misses his final end” (Anth 7:271).

By self-deceptively exaggerating the value of some imagined object, a passion makes it difficult for us to consider rationally the strength of the incentive it offers us. The subject of a passion is often willing to run extreme risks, or even engage compulsively in directly counterproductive behavior, in pursuit of the object. An ambitious person, for instance, who seeks honor from others wants the love and approval of others and the pleasures of social intercourse with them. “However, if he is a passionately ambitious person, then he is blind to these ends, though his inclinations still summon him to them, and he overlooks completely the risk he is running that he will be hated by others, or avoided in social intercourse, or impoverished through his expenditures” (Anth 7:266). Passions can therefore be called “inclinations of delusion, which consists in valuing the mere opinion of others regarding the worth of things as equal to their real worth” (Anth 7: 270). We see here that not only the object but even the delusion involved in passions is social in context and content. Where morality is at stake, the self-devised delusions involved in passion play an essential role in the irrational preference of incentives of inclination over those of morality. Passions are therefore always evil and conspicuous manifestations of the propensity to evil in human nature (VA 7:267). That people are prone to inflict such delusions on themselves is, of course, fundamentally just as incomprehensible as the choice of the evil maxim itself. But it clearly plays a role in that choice.

Kant divides passions into “natural” and “social.” Natural passions, every bit as much as social passions, are directed at other human beings, but they

belong to what is “innate,” and not from *culture*, or what is *acquired* (Anth 7:267). Kant identifies two natural passions: for *freedom* and for *sex* (Anth 7:268, cf. VA-Mrongovius 25:1359–1361), though he also discusses the passion for vengeance under the heading of natural passions (Anth 7:270–271).

Natural passions. “Freedom” is of course a term with many senses, and in Kant’s philosophy it often refers to something (or many things) that are good or at least that have good aspects. But Kant regards *freedom* (external freedom, independence of the will of another) as the most fundamental object of human inclinations. “The most terrible condition for a human being is that in which another human being determines his condition, and cares for his happiness in accordance with his inclinations” (VA-Collins 25:214, cf. VA-Friedländer 25:581). “Whoever is able to be happy only according to another person’s choice (no matter how benevolent this other person may be) rightly feels that he is unhappy” (Anth 7:268).

For the same reason, freedom is the object of the most violent of all human passions. Kant thinks that the cry of the newborn infant is a wail of rage, occasioned by its inability to use its limbs, which it experiences as a kind of external constraint (Anth 7:268). According to Kant the brutal savagery of “uncivilized” peoples is due to their refusal to submit to the external rule of right, which is a condition for all agriculture and the civilized society based upon it (Anth 7:269, cf. R 6:32–33). “The savage (who is not habituated to submission) knows no greater misfortune than falling into [the power of another]; and he is right, as long as there is no public law to protect him ... This accounts for his constant state of warfare, by which he intends to keep others as far away from him as possible” (Anth 7:268). Here “we see in the complacency with which the victors praise their great deeds (massacring whole peoples, swooping down on them without sparing any, and the like) that what they really value in these deeds is only their own superiority and the destruction they are able to wreak” (R 6:33). The passion for freedom is thus only the “uncivilized” form taken by the social passion for domination over others: “The human being’s self-will is always ready to break forth into hostility toward his neighbors, and always presses him to claim unconditional freedom, not merely independence of others but even mastery over other beings that are his equals by nature” (Anth 7:327).

Kant understands sexual passion as the desire to use another person’s body for the gratification of merely animal impulses (Anth 7:136). Because it is directed not at the whole person but only at the person’s sex organs, and through the use of this part (without considering the person as a whole), it reduces the person to only this physical part and treats the person as a mere object or thing. The meaning of sexual desire, therefore, is to dominate and degrade. “Because sexuality is not an inclination which one human being has for another as such but is an inclination for the sex of the other, it is a principle of the degradation of humanity, ... a preference for dishonoring [humanity] in order to satisfy an inclination” (VMO-Collins 27:385).

Kant allows that sexual desire can be combined with genuine love, though he regards the combination as inherently unstable. He also sees sexual desire, imaginatively transformed and combined in the social condition with the experience of sexual *refusal*, as the psychosocial source of *respect* for personality, of moral feelings generally, and of the human capacity to control inclinations through reason (MA 8:113). It is a striking and underappreciated aspect of Kant's moral psychology that he thinks *the fundamental moral feeling of respect is rooted in human sexuality* (for further discussion, see Wood, 2008, chap. 13). But sexual *passion* is not the source of morality or respect, except in the sense that it is the form of human desire that makes respect necessary and poses a constant threat to it. Kant's views on sexuality and on women contain a lot that is nowadays either laughable or morally repugnant to enlightened sensibilities. But this should not be allowed to conceal from us the fact that one of Kant's main preoccupations in this area is with protecting the woman's basic right and dignity from assaults (economic and social as well as direct and violent) that are made possible by the physical and social superiority that men have over women (again, see Wood, 2008, chap. 13). There is nothing outdated or unenlightened about that concern. It is as up to date as the latest headlines.

Social passions. In the *Groundwork* Kant lists among the "gifts of fortune" the trio of power, wealth, and honor (G 4:393). We probably do not appreciate that here he is also giving us, in outline, and by way of their respective objects, his theory of the social passions: *tyranny* (or the addiction to domination, *Herrschaftsucht*), *greed* (the addiction to possessing, *Habsucht*), and *ambition* (the addiction to honor, *Ehrensucht*).

The same trio lies behind Kant's earlier threefold division of "resources" (as formal objects of inclination): "Resources are threefold: talent, power and money; on this are grounded the three inclinations: ambition, tyranny and greed. These are the three passions, which pertain to the three resources through which we seek to satisfy all our inclinations" (VA-*Menschenkunde* 25:1141). But the general aim of the social passions consists of a pair of related objects: "obtaining rank or status among [our] fellow human beings" (Idea 8:21) and "getting other people's inclinations into one's power, so that one can direct and determine them according to one's intentions, [which is] almost the same as possessing others as mere tools of one's will" (Anth 7:271).

3.9 THE THREE DEGREES OF EVIL

Human self-deception is endlessly inventive and manifests itself in more forms than anyone could possibly list. But the fundamental form of self-deception that lies at the ground of all evil is a lie to ourselves about *who we are*. What each of us fundamentally is, is a self-governing rational agent belonging to a community of rational agents who are equally ends in themselves. Whatever other value we may attach to ourselves is dependent on that. This identity makes certain

rational principles – instrumental, prudential, and ultimately moral – binding on us and gives the moral incentive rational priority to all others. This is not something we can fail to be aware of if we truly are rational agents, and the choice not to act according to who we are is possible only through a self-deceptive delusion that some incentive other than the moral one has rational priority over it.

Kant conceptualizes the propensity to radical evil in the form of a highly abstract *maxim*: the maxim of preferring the rationally inferior incentives of inclination and self-love to the rationally stronger incentive of morality. Despite its simplicity and abstractness, the manifestations of this maxim are supposed to be as luxuriantly diverse and inventive as humans are in manifesting evil. Kant tries to give us a sense for this subtlety and creative opulence by distinguishing three “degrees” in which the propensity to evil may show itself. Though they can be manifested in actions, as degrees of the *propensity* to evil they are degrees of *temptation*.

These three degrees of evil are:

- (1) “frailty” (*Gebrechlichkeit, fragilitas*),
- (2) “impurity” (*Unlauterkeit, impuritas, improbitas*), and
- (3) “depravity” or “perversity” (*Bösartigkeit, Verderbtheit, vitiositas, pravitas, perversitas, corruptio*) (R 6:29–30).

The lowest degree of evil is what Kant calls the “frailty” of the human will that adopts a good maxim but proves too morally weak (too lacking in virtue) to carry it out. The “frail” will deceives itself insofar as it represents its failure to live up to its good maxim as a casual, perhaps merely “accidental” failure, a temporary “lapse,” whereas in fact it could not have occurred but for a systematic choice to allow the rationally inferior incentives of self-love or inclination to take precedence over those of reason or the moral law.

The second degree of evil is “impurity.” It is noteworthy that impurity represents a greater temptation to evil than frailty because frailty shows itself in evil action, while impurity need not. With it, “although the maxim is good with respect to its object (the intended compliance with the law) and perhaps even powerful enough in practice, it is not purely moral, i.e. [has not] as it should have, adopted the law alone as its sufficient incentive, but, on the contrary, often (and perhaps always) needs still other incentives besides it in order to determine the power of choice for what duty requires” (R 6:30). What Kant means by “impurity” is *not* the case Kant considers early in the *Groundwork* where the mere presence of nonmoral incentives for doing a dutiful action makes it the case that there is no need for (hence also no possibility of) acting “from duty” (G 4:397–399). Rather, impurity occurs not only where nonmoral incentives are present but where there is a systematic *dependence* on them, a weakness of character where the nonmoral incentives are self-deceptively confused with the incentive of duty. Impurity misleads us into evil action chiefly by allowing inclinations to “enter into the determination

of duty” – that is, into corrupt and mistaken judgments about what it is our duty to do (TP 8: 279). In Kant’s formulas of the moral law, it is the function of the Formulas of Universal Law and Law of Nature to provide a “standard” or “canon” of judgment that can help us correct such culpable errors (G 4:402–405, 421–425). (For the argument that these formulas have no other function, see Wood, 2017, pp. 13–60.)

The highest degree of temptation is displayed in “depravity” or “perversity.” Here there might seem to be no room for deception of any kind because depravity appears to be only the brazen and open choice of “maxims that subordinate the incentives of the moral law to others (not moral ones)” (R 6:30). But we must not ignore the essential point that the concept of evil makes sense only if morality carries reasons with it that are overriding in relation to other reasons, such as those of self-love or inclination. Because the rational will cannot be “diabolical,” evil can never be anything but a failure of rational volition. And there would be no genuine evil if the supposedly evil agent chose other incentives over those of morality only through an innocent error about their relative rational strength – as if, for instance, someone simply had no way of knowing that the rights of others give him good reasons to limit his pursuit of the pleasure he takes in harming them. An agent who was truly ignorant about that could not be a rational or moral agent at all.

The concept of evil makes sense at all only if the one who chooses evil is well aware that the incentives of morality are rationally superior to nonmoral incentives but gives preference to the latter anyway. Such a deliberately perverse choice, however open, brazen, and shameless it may represent itself as being on the surface, or whatever rationalizations it may use to cloak itself, necessarily involves a deep self-deception on the agent’s part about the structure of its own commitment to what it knows about the relative rational force of the incentives by which it is moved. At the denouement of a melodrama, when the villain is caught in the act, he sometimes brazenly declares that he and he alone is rational and honest with himself, while his accusers are all cowards, liars, and hypocrites. This nasty little speech may sometimes be accurate in its depiction of the accusers, but as regards the villain himself, it necessarily represents the uttermost extreme in desperate self-deception. All those who find it credible are badly off, both psychologically and morally.

3.10 USE OF THE THESIS OF RADICAL EVIL

Regarding the thesis in part one of the *Religion* that human nature contains a universal, innate, and inextirpable propensity to evil, the literature contains a remarkable schizophrenia: one set of commentators charges that Kant’s thesis is so extreme as to be incoherent, even self-contradictory. Another set claims that the thesis is true a priori and either that Kant himself has provided an a priori proof of it or that one can be provided on his behalf. Both of these sets of extreme claims seem to me clearly mistaken. Kant’s thesis is indeed very strong,

far too strong for it to be proven a priori. But it is carefully crafted not to be self-contradictory.

Some critics have charged that Kant is ultimately faced with a contradiction he cannot escape when he claims that evil is both imputable and also present in human nature. Some worry about the claim that it is “innate,” others about the claim that it is “inextirpable.”²⁴ As we have already seen, Kant’s claim that the propensity is “innate” refers not to its causality (which Kant holds is “inscrutable”) but to there being no identifiable temporal point in our lives when it has not already been acquired (R 6:38). It is not contradictory to say that we are responsible for something even though we cannot point to a specific time when we chose it. People are blamed for their hatred, bigotry, greed, callousness, and other bad traits of character even though we cannot identify a precise time at which they acquired these vicious traits. Kant does not consider these traits “innate,” but that is presumably because, unlike the propensity to evil (which is their root), they are not fundamental (or “radical”), hence not ubiquitous. Nor is there anything contradictory about holding that we can freely and imputably do something to ourselves that we cannot then undo. It belongs to time’s arrow that some things once done cannot be undone: spilt milk cannot be put back into the bottle; toothpaste cannot be squeezed back into the tube. It is indeed an extremely strong claim that all human beings, without exception, have inflicted on themselves a propensity to evil, especially that at every point in their lives all human beings have *already* done so. But it is not self-contradictory.

It is implausible beyond imagining that so extreme a thesis as Kant’s thesis of the radical evil in human nature could be proven a priori. Above in §5 we have discussed the sole passage that could be read (i.e., misread) as suggesting such a thing (R 6:20–21). There is, however, a sizable literature that tries to do just that, or claims that Kant tried to do it.²⁵ There is no space here to show how each of these attempts fails, how each misconstrues which passages in Kant, and what elements of insight each may display along with the fundamental implausibility of what each is trying to argue.

I will content myself with quoting Kant’s own clear declaration concerning what he thinks it would take to establish the thesis:

However, that by the “human being” of whom we say that he is good or evil by nature we are entitled to understand not individuals (for otherwise one human being could be assumed to be by nature good and another evil), but the whole species – this can be

²⁴ Innateness is what worries Michalson, 1990, pp. 46–66. Inextirpability is the problem for Bernstein, 2002, pp. 11–45.

²⁵ See Allison, 1990, pp. 152–61; Palmquist, 2000, pp. 156–58; Palmquist, 2008, pp. 261–97; Morgan, 2005, pp. 63–114; and Muchnik, 2010, pp. 116–143. Papish, 2018, p. 118 and chap. 8, claims that I regard unsociable sociability as part of a “strategy” to “prove” the thesis of radical evil. But I deny there is any “proof.” Kant says explicitly at R 6:93–95 that empirical evidence for unsociable sociability makes the thesis of radical evil more plausible. I say no more than that.

proven (*bewiesen*) only later on (*weiterhin*) if it is shown in anthropological research that the grounds that justify us in attributing one of these two characters to a human being as innate are of such a nature that there is no cause for exempting anyone from it, and that the character therefore applies to the species (R 6:25–26).

Here it seems quite evident that the “proof” Kant has in mind is *not* a transcendental deduction or an a priori demonstration of any kind. It rather consists in the results of “anthropological research” – and indeed a research project that would have to be completed “later on.” The “formality” of this proof would presumably consist in a proper scientific presentation of the grounds (partly rational and partly empirical, as in any natural science conceived in Kantian terms) for the conclusion that the evil propensity belongs innately to all members of the human species. We are given the methodological hints toward such a proof in the teleological conjectures Kant presents in *Idea for a Universal History*, where it is argued that nature provides for the full development of human predispositions through the mechanism of unsociable sociability (I 8:21–22). The actual proof would be provided by a future empirical anthropologist-historian who successfully followed Kant’s philosophical idea (I 8:29–31). But these are only methodological hints. What is unmistakably clear from the passage just quoted is that Kant thinks the thesis, and especially the universal attribution of a propensity to evil to all human beings, has *not yet been proven* at all, much less proven a priori. This ought to make us wonder about how far, and for what purposes, Kant himself is even committed to the thesis of radical evil.

Section III of part one is boldly titled “The Human Being Is by Nature Evil” (R 6:32–39). Here Kant asserts that “we may spare ourselves the formal proof (*förmlichen Beweis*) that there must be such a corrupt propensity rooted in the human being, in view of the multitude of woeful examples that the experience of human deeds parades before us” (R 6:33). He then offers us an all too familiar account of the notorious barbaric cruelty of so-called natural or primitive peoples and what he regards as the even worse vices of civilization – falsity, hypocrisy, treachery to friends, ingratitude, the malicious pleasure sophisticated people take in the misfortunes even of those closest to them (R 6:33–34). This is obviously not an argument from induction or an empirical generalization from random instances. The idea is to consider humanity in its most extreme contrasting forms – the “primitive” and the “civilized” – and perhaps these observations do make it more plausible to consider the propensity to evil universal.

Kant clearly realizes that it falls short of establishing the thesis of radical evil. One might even have expected this given how extreme the thesis is. If its “formal proof” must wait upon the results of “anthropological research later on” then Kant is saying that such a proof does not yet exist. Writing for an audience of faithful Lutherans, who presumably accept the Christian doctrine of Original Sin, he is not eager to stress this point; but neither does he attempt to conceal it.²⁶

²⁶ Some scholars find such an honest confession of so obvious a point shocking or even scandalous.

“Kant never gives – or even attempts to give – a proof of his controversial and bold claim that

Kant begins part one of the *Religion* by putting before us the proposition that “the world lieth in evil” (R 6:18; cf. John 5:19). As we have seen, some are convinced that Kant has offered an a priori proof of radical evil, while others are shocked that he fails to provide any proof for it. But Kant poses this, as a disputed proposition, opposed by the “heroic” and “modern” view that the world “forges steadfastly (though hardly noticeably) ahead ... from bad to better” (R 6:20). A careful reader should be able to see that Kant wants to agree in at least a qualified way with both sides of the dispute. He presents sympathetically the thesis of radical evil but without dogmatically accepting it. In fact, his sympathies lie at least as much on the “modern” and “heroic” side.

Several years later Kant’s *Anthropology* concludes its attempt to describe the “character of the human species” with these words:

If one now asks whether the human species ... is to be regarded as a good or bad race ... [then] it presents [itself] not as evil, but as a species of rational beings that strives among obstacles to rise out of evil in constant progress toward the good. In this its volition is generally good, but achievement is difficult because one cannot expect to reach the good by the free agreement of *individuals*, but only by a progressive organization of citizens of the earth into and toward the species as a system that is cosmopolitically united (Anth 7:331–333).

This passage does not explicitly reject the thesis that human nature contains a universal, radical, innate, and inextirpable propensity to evil. But it certainly does not endorse it either. We know from Kant’s other writings that he thinks it cannot be proven that the human species is morally improving but also thinks that we have a moral obligation to contribute to its improvement and that this gives us a practical ground for assenting to the thesis that it is morally improving (TP 8:307–313). This is clearly the sentiment he is expressing at the end of the *Anthropology*. Kant even thinks we are licensed to look cautiously for empirical evidence of the moral improvement of humanity and that we can find it in the form of the reaction of enlightened spectators of the recent political revolution in France (SF 7:79–94).

In light of this, we should not ask, How does Kant prove the thesis of radical evil? For he does not even try or claim it has yet been “proven” at all. Kant regards the thesis of radical evil as plausible, on empirical grounds, but not demonstrable, either empirically or a priori. The question should be rather, Can this plausible but dubious thesis of radical evil find a place in a pure religion of reason, and if so, how? Kant provides an explicit answer in the General Remark at the end of part one:

man is evil by nature”; they are appalled by the “absence of genuine argumentation on this crucial point” (Michalson, 1990, p. 46; cf. Bernstein, 2002, pp. 34–35). Goethe was famously shocked at Kant’s apparent endorsement of the Christian doctrine of original sin, while Michalson and Bernstein are shocked that he does not offer conclusive proof of it.

The moral law [however], commands that we *ought* to be better human beings now, and it inescapably follows from this that we must be capable of being better human beings. The thesis of innate evil is of no use in moral *dogmatics*, for the precepts of the latter would include the very same duties, and retain the same force, whether there is in us an innate propensity to transgression or not. In moral *discipline*, however, the thesis means more, yet not more than this: We cannot start out in the ethical training of our connatural moral predisposition to the good with an innocence which is natural to us but must rather begin from the presupposition of a depravity of our power of choice in adopting maxims contrary to the original ethical predisposition, and since this propensity is inextirpable, with unremitting counteraction against it (R 6:50–51).

Kant then goes on to argue that for the purposes of *moral discipline* (i.e., for the way people go about trying to improve themselves morally), the thesis of radical evil entails that improvement not only depends on a change of deeds or habits but requires as the ground of all these a moral conversion, a radical revolution or *change of heart* (*Herzensänderung*). This change of heart is a moral duty and the sole ground of our hope that we may become well-pleasing to God. This *consequence* of the thesis of radical evil for moral practice is what is essential to the overall argument of the *Religion*. It will be the topic of Chapter 4.

The Change of Heart

In Chapter 3, we saw that Kant treats the thesis of radical evil as not yet proven but of use for moral “discipline”: we cannot begin our strivings for moral improvement with a presumption of our innocence, but must presuppose an original depravity of our disposition, and therefore we need to combat it (R 6:50–51). The first consequence of the thesis of radical evil for moral discipline must be that we restore the rational order of incentives, adopting the maxim of putting the moral incentive ahead of all nonmoral ones. Kant calls the radical decision to do this the “moral revolution” or “change of heart.” It is grasped symbolically, in Christian terms, as a “rebirth,” even a “putting off of the *old man*” and a “putting on of the *new man*.”¹

Under the influence of the propensity to evil, the human being has an “evil heart” (R 6:29). The moral revolution is the fundamental change from an evil heart to a good heart. But if the propensity to evil is not only “innate” but also “inextirpable,” then nothing can effect an immediate and complete change from evil to good. The change of heart consists in a change in “attitude of mind” or “way of thinking” (*Denkungsart*). But this is not literally a single sudden change at a time, though we can represent it to ourselves symbolically as one. What it can be in time is only a gradual, open-ended struggle for moral improvement in our future course of life:

¹ The German terms “alter Mensch” and “neuer Mensch” are drawn from the Lutheran New Testament. Though the German noun “Mensch” is grammatically masculine, it would be more accurately rendered “old human being” and “new human being.” But if we are to preserve the allusion to the Bible, in the King James and most other English versions, the masculine noun is necessary. The same is obviously true of the “Son of God,” which is our theme in the next couple of chapters. It is of course also implicit in the Christian symbolism that the referents of both “old human being” and “new human being” are also to male human beings: Adam and Jesus Christ. This note is intended to explain my translation practice and to acknowledge how problematic it always has been to treat male human beings as paradigmatic representatives of the entire human species. It is problematic not only for Kant and the entire Christian tradition but for others too.

If by a single and unalterable decision a human being reverses the supreme ground of his maxims by which he was an evil human being (and thereby puts on “a new man”), he is to this extent, by principle and attitude of mind, a subject receptive to the good; but he is a good human being only in incessant laboring and becoming, i.e. he can hope . . . to find himself upon the good (though narrow) path of constant progress from bad to better. For him who penetrates to the intelligible ground of the heart . . . this endless progress is a unity, i.e. for God; and to this extent the change can be considered a revolution (R 6:48).

4.1 THE PURITY OF THE MORAL INCENTIVE

The fundamental claim Kant is making here is already a familiar part of his own moral philosophy. Kant holds that there is an a priori supreme principle of morality. He calls the moral principle a priori because our reason gives the moral law to itself through its own activity rather than receiving it passively through external influences. Of course we must be educated to become rational beings through empirical interaction with objects and especially with other human beings (VP 9:441). Reason, however, is not a mechanical product of conditioning but an active faculty giving itself its own moral principle. This is why Kant calls the moral law “a priori” and says that “pure reason is practical of itself alone” (KpV 5:31).

What is perhaps most striking in Kant’s moral philosophy, however, is that the a priori moral law contains a unique reason, *motive* (*Bewegungsgrund*) (G 4:427) or *incentive* (*Triebfeder*) (KpV 5:71–89), distinct from all those offered us by empirical impulses, desires, and inclinations. Kant’s argument is based on the premise that *necessarily*, dutiful actions are valuable for their own sake and not merely because they produce independently desirable results. We need experience to know which actions conform to duty, by applying the a priori moral principle. But the pure rational incentive to do these actions depends neither on our intuitions about what is good nor on our contingent, empirical attitudes toward these actions. This reason is fundamentally unitary – a single basic reason grounded a priori solely in our rational faculty.

Kant is convinced that the distinctiveness of the moral incentive matters not merely for moral philosophy but for moral practice itself – for the lives of ordinary moral agents:

Morals themselves remain subject to all sorts of corruption as long as that guiding thread and supreme norm of their correct judgment is lacking. For as to what is to be morally good, it is not enough that it *conform* to the moral law, but it must also happen *for the sake of this law*; otherwise, that conformity is only contingent and precarious, because the unmoral ground will now and then produce lawful actions, but more often actions contrary to the law (G 4:390).

Kant thinks, moreover, that only a moral consciousness focused on the distinctive moral incentive can contribute to true *moral improvement* in the

world. In the *Groundwork* Kant cites a letter he received from the late J. G. Sulzer (1720–1779) that he interprets as raising the question why moral theories, though convincing to reason, accomplish so little toward making people morally better. Kant’s answer is that the failure lies in the fact that these theories do not represent morality in its purity – that is, they do not concentrate attention on the one distinctive rational incentive that is essential to morality. Instead they appeal to a variety of different motives that are only contingently, hence precariously, related to moral conduct (G 4:411 n, cf. 4:390). Just before this, Kant has been criticizing popular moral philosophy on these grounds, complaining about the variety of incentives it offers on behalf of morality: “sometimes perfection, sometimes happiness, here moral feeling, there fear of God, some of this and some of that, all in a wondrous mixture” (G 4:410). Such philosophers think they are helping the cause of morality by heaping up the variety of attractions associated with moral conduct. But, Kant protests, “in trying to strengthen their medicine they ruin it.” In the second Critique, Kant begins his argument by declaring that “material principles” – those that draw their incentives from the desirability of ends to be produced – are all of the same kind: not only are they distinct from the moral principle but they have a corrupting influence on morality (KpV 5:19–29, 82–89).

Those who fail to recognize and value the distinctiveness of the moral incentive, or who mix and confuse it with other incentives, are not merely embracing an erroneous philosophical theory. Their error makes them *morally worse people*. This is because recognition of the primacy of the moral incentive is equivalent to clearheadedness (absence of self-deception) about my fundamental identity as a free, self-governing, rational moral agent. To substitute any other identity for this one, or to mix the incentives belonging to this identity with competing incentives belonging to a false and competing identity, is already to manifest the radical propensity to evil.

4.2 GOODNESS OF HEART

What is this distinctive moral incentive? Kant’s presentation in his foundational works of the distinctive moral incentive tends to emphasize the thoughts of *duty* and the *moral law*. To many, the associations attaching to these terms – the implication that it is cold, abstract, even legalistic – are largely responsible for most of the attitudes of visceral rejection with which Kant’s moral theory is often greeted. These attitudes all rest on a common but deplorable error. Kantian ethics is an ethics of sympathetic participation (*Teilnehmung*) and caring. Kant’s moral psychology holds that reason grounds feelings, toward ourselves, our actions, other human beings, other living things and nature (MS 6:399–403), without which we would not be rational agents at all. No doubt in some contexts doing the right thing requires self-constraint, and Kant thinks these contexts tell us most about what morality is. But for Kant moral reason

itself is not cold or legalistic. For it is another, equally mistaken popular prejudice that reason is opposed to emotion. Emotions can be irrational, but they are always to some degree based on reason; and for Kant reason must express itself in feeling and emotion. We should associate Kant's incentive of "duty" (*Pflicht*) with the verb *pflegen* (to care for), which is the etymological source of the term.

In the *Groundwork* and second Critique, Kant associates the moral incentive with the *good will*, and names it only "respect for the moral law" (G 4:400–402, KpV 5:73–79). In the Doctrine of Virtue, however, we see that the moral incentive consists in the basic *feelings* of *love* and *respect* toward others (MS 6:401–403, 448–450). In the *Religion* Kant calls giving first priority to the moral incentive having a *good heart* (R 6:29, 50, cf. G 4:410, KpV 5:156, MS 6:441). Goodness of heart is characterized by the feelings of love and respect; they are what people most need from one another; the giving and the receiving of them is the only thing that makes human life worth living. Love and respect must, however, be kept in balance. For love without respect always tends toward condescension and ends in contempt; respect without love tends toward coldness, distance, resentment, and finally hatred. The moral incentive and goodness of heart consist in this proper balance of love and respect. For beings like us, this does not come easily. Maintaining this balance and bringing a good heart to bear on our actions are often something we need to constrain ourselves to do. So *duty*, as rational self-constraint, is morally indispensable.

The moral incentive is the ground of a good heart. Kant's conception of moral virtue, however, also gives *empirical* feelings and inclinations a significant role. Kantian ethics mistrusts empirical incentives when they are divorced from moral reason and are not expressions of a good heart. But empirical desires, feelings, and habits are the necessary *consequences* of a *virtuous disposition*; in fact, they are the only true *signs* by which a virtuous disposition can show itself in conduct. The association of moral conduct with empirical feelings and desires is precisely what constitutes the appearance of virtue: *virtus phaenomenon*.

When the firm resolve to comply with one's duty has become a habit, it is called *virtue* in a legal sense, in its empirical character (*virtus phaenomenon*). Virtue here has the abiding maxim of lawful actions, no matter whence one draws the incentives that the power of choice needs for such actions. Virtue in this sense is accordingly acquired *little by little*, and to some it means long habituation [as] a human being, through gradual reformation of conduct and consolidation of his maxims, passes from a propensity to vice to its opposite (R 6:47).

Virtus phaenomenon – the *appearance of virtue* – refers simultaneously to two different and even apparently opposite things, which are, however, also inseparably connected: first, it means *apparent virtue*, which may be a deceptive semblance and not true virtue at all; second, it means the *manifestation of virtue*, the only way that true virtue could ever show itself in our conduct.

In Kant's moral psychology, inclinations are not causal levers operating mechanically on the agent. That conception of them would be incompatible with freedom of the will, and it would also render true virtue impossible. Inclinations are never *causes*; they are grounds, or reasons. This Kantian doctrine is emphasized in what Henry Allison has called "the incorporation thesis." Inclinations can play a role in our actions only when they have been incorporated into a maxim freely chosen by us (R 6:24). But Kant's views about this matter are even more complex. When we act on inclinations independently of the moral incentive, giving them priority over it, they are expressions of the evil propensity and its evil maxim. In an agent with a good disposition, however, the right empirical feelings and desires become reasons expressive of the good maxim or pure disposition, which gives priority to the moral incentive. When this happens, *virtus phaenomenon* is an expression of *virtus noumenon*, the moral incentive, true goodness of heart. In our conduct of life, this unavoidable ambiguity is a persistent reality that should worry us.

Neither by introspection nor by observing our external conduct can we know whether our dutiful conduct and our duty-favorable inclinations are expressions of a good and pure disposition. But that disposition alone is what distinguishes the true change of heart from anything that we could directly observe in ourselves or others. The acquisition of virtue in the form of steady, gradual improvement in our conduct, and in the feelings and habits that go along with good conduct and support it, turns out to be the *only possible evidence* that the change of heart has actually occurred. "For a human being who, from the time of his adoption of the principles of the good and throughout a sufficiently long life henceforth has perceived the efficacy of these principles on what he does, has cause to infer, but only by way of conjecture, a fundamental improvement in his disposition" (R 6:68). From our standpoint, and from that of any temporal being who sees our conduct only in part and not in its completeness, this *virtus phaenomenon* might be only a false *semblance* of virtue; but it might also be the only possible *manifestation* of *true* virtue. Whether it is genuine virtue or mere semblance depends on the whole history of our conduct and whether it exhibits the gradual progress that is the only possible manifestation of a genuine change of heart. Thus Kant's quarrel with the popular moralists (G 4:407) is that they mistake certain appearances or signs of virtue for what is essential to it. He conceptualizes true virtue by saying it results from the single decisive change of heart that revolutionizes our disposition, putting us on the path to good.

Can we know during our life whether the change of heart has actually occurred, when *virtus phaenomenon* is genuine virtue and when not? No. This is something we can never know, not even in principle. As long as we live and act: during our lives in time, there is not even a determinate fact about it. Our conduct always leaves it ambiguous. It is fatally misleading to represent this situation, as Kant's readers are often tempted to do, as if this change were something already actual but hidden from us somewhere else – in a timeless

noumenal world intellectually intuited by God. Kant thinks we suppose God to cognize our disposition through intellectual intuition but not by perceiving hidden psychic mechanisms within us and certainly not by cognition of something belonging only to an ectoplasmic self in a supernatural world. God's knowledge of our disposition consists only in the complete knowledge and wise judgment of the whole course of our temporal life, for it is there alone that our true disposition could ever be manifested.

4.3 ACTING FROM AN INCENTIVE

In his account of the change of heart, Kant lays a great deal of stress on the idea that it is a decision to "become not merely legally good, but morally good (pleasing to God) ... and thus in need of no other incentive to recognize a duty except the representation of duty itself" (R 6:47). This means that the Kantian distinction between *legality* and *morality* is central to understanding the change of heart. Kant also equates this distinction, in biblical terms, with the Pauline distinction between acting according to the *letter* of the law and acting according to its *spirit* (R 6:30, 195; cf. 2 Cor 3:6). The distinction between legality and morality is closely connected to a difficult set of Kantian ideas: acting from duty (*aus Pflicht*), as distinct from merely acting in conformity with duty (*pflichtmäßig*) (G 4:397–399); acting for duty's sake or for the sake of the law (*um des Gesetzes willen*) (G 4:390); and fulfilling the *duty* to act *from duty* (MS 6:391, 393).

There is clearly a central theme here, but it is all too common to form oversimplified and mistaken ideas about it, which we are bound to do if we assume a moral psychology alien and opposed to Kant's. Consider this deceptively simple formulation:

What is essential to any moral worth of actions is *that the moral law determine the will immediately*. If the determination of the will takes place conformably to the moral law but only by means of a feeling, of whatever kind that has to be presupposed in order for the law to become a sufficient incentive, so that the action is not done *for the sake of the law*, then the action will contain *legality* but not *morality* (KpV 5:71).

The first mistake one is likely to make in reading this passage is that of thinking that "determining the will" refers to *causal* determination. That is, we might be tempted to suppose that each individual act has a determinate psychic cause. To act with *morality*, for the sake of the law or from duty, is to have one's action psychically *caused* by the incentive of duty. An action with mere *legality* conforms to duty but is *caused* by a different incentive: one of inclination, self-love, or empirical feeling. There are indeed some great philosophers – Spinoza, Hume, perhaps also Schopenhauer, and many since – who do think of our actions as "motivated," meaning simply that they are *caused*. For them reasons and motives are causes, and the causes of actions are found either in external events that causally influence us or in our mental processes or in both. The

causal approach to motivation is also the approach often taken by empirical researchers when they investigate “moral motivation.” That motives are causes is not an empirical conclusion confirmed by their research; it is an unquestioned assumption tacitly built into it and taken for granted in the very questions it asks, the very hypotheses it is able to entertain. For these philosophers and psychologists, “motivation” is all about the *aetiology* of action.²

It is important to appreciate that Kant is *not* one of these philosophers. Kant does think of the free will or power of choice (*Willkür*) as a *cause of what results* from the action, even a cause that has its own law – the moral law (G 4:446–447). And the *existence* of a finite freely willing being is causally dependent on many things. But Kant holds that free volition itself is not a *caused* cause; it is a free or uncaused cause, or rather a *self-caused* cause. The free will is a cause that acts entirely from itself or oneself (*von selbst*) (KrV A 533/B 561). For Kant the concept of cause contains *necessitation*: when something is *caused* to happen, once the cause is posited, the effect cannot happen otherwise than it does. This cannot be true, however, of our free volitions. Kant thinks that for positively free beings like ourselves, all actions not in conformity with duty could have been otherwise, and for imperfect beings like ourselves, self-afflicted with a propensity to evil, every action in conformity with duty could have been otherwise – a wrong action. Consequently, none of our *free* volitions can be *causally* determined (naturally necessitated) by anything whatever.

Thus when Kant speaks of the “determination of the will” (*Willensbestimmung*), we must understand the verb *bestimmen* in a different sense: the sense in which any subject is *determined* by a predicate that in fact applies to it (KrV A 571–573/B 599–600). When Kant says the will is “determined” by practical reason, or, alternatively, by some other incentive, he should be understood to say that the will is *characterized* by whatever incentive provides the ground or reason for its freely chosen action.

But what is it for an incentive to “determine” (to characterize) a volition or action? The wrong answer to that question will lead to a different but equally disastrous misunderstanding. Does Kant think that in freely choosing each action, we also at the same time choose the incentive from which that particular action is done, so that it thereby acquires the predicate that it is done from that incentive? No, he does not. For if he did, then it would follow that the duty to act from duty would be simply a strict or narrow duty – namely, the duty to choose, along with each dutiful action, also the incentive of duty as that from which it is done.³ Every action in conformity with duty would then

² For one example, see Schroeder, Roskies, and Nichols, in Doris (ed.), 2010, chap. 3.

³ That invites a *reductio ad absurdum* argument proposed by H. A. Prichard and W. D. Ross (Prichard, 2002, pp. 3, 11; Ross, 2002, pp. 5–6): suppose there is a perfect duty that every obligatory action, e.g., action A, must be done from the motive of duty. In that case, doing A from the motive of duty would be yet another obligatory act (call it A'). A' is then also a narrow or perfect duty. And it would of course follow that there is also a perfect duty to do A' from duty, which makes into a duty yet a third act: A'': doing A' from duty. And so on ad infinitum ... Action

involve a transgression of duty if the incentive chosen were anything but the pure incentive of duty. Even if the dutiful action itself fulfills an imperfect duty, the duty to choose the incentive of duty would then have to be a perfect duty because to choose any other incentive would be to choose against the moral law. That is not Kant's position. He holds that although the duty to act from duty might look like a perfect duty, in fact, it is not; it is a wide or imperfect duty, even in the case of actions that are themselves perfect duties (MS 6:392). Kant says that the duty to act from duty is a "*narrow* and perfect one in terms of its quality, but it is wide and imperfect in terms of its degree, because of the frailty [*fragilitas*] of human nature" (MS 6:446). We will return later to this puzzling claim. His conception of the change of heart in the *Religion* will help us to understand it.

If we are to grasp Kant's position, we need to think of "being motivated" or "acting from an incentive" in a different way. Kant is closer to Elizabeth Anscombe, who holds that to attribute a motive to an action is "to put it in a certain light," or to Hegel, who considers motivation a way of bringing an action under a certain abstract concept that indicates the agent's reason for doing it (Anscombe, 1963, p. 21; Hegel, 1991, § 119).⁴ An incentive or reason is *a way of thinking* about the action when you freely choose it. You are motivated to do A by incentive T if in doing A, you think about A in light of T in your choice to do it. This *way of thinking*, moreover, is not an occurrent state (either of being caused or of choosing) that accompanies the particular action. It is not something "operative" here and now in the process of choosing. It is instead a disposition, expressing your character or the kind of person you are. But because you are free, and in time, it is always essentially ambiguous, hence hidden – from you, from all of us – not hidden away in another ("noumenal") world but hidden from us right here in this one. Your disposition or character also depends on your choices but not on any single occurrent choice. It is what you have made of yourself, which is related in complex ways to what you have done and also to what you are disposed to do (Anth 7:291–292). It is a matter of *rational* habituation.⁵

Kant never considers a case where an agent has both the incentive of duty and a nonmoral incentive for the same dutiful action and is then said to act "from duty" because the former incentive rather than the latter is "operative" in that

from duty would require an infinite regress of dutiful acts, hence be impossible. But Kant subscribes to none of this nonsense. He denies the duty to act from duty is a perfect duty; it is *not* a duty to perform the inner action of choosing the incentive of your outer action (MS 6:393).

⁴ Donald Davidson (1963) argues that if there is to be a determinate fact about the reason (or incentive) from which we do an individual action, then reasons must be a kind of cause. Kant would accept that premise but use it contrapositively: since reasons (or incentives) cannot be causes, it follows, as we have already seen, that for any specific action there can be no unambiguous fact that it is done from any specific incentive. Motivation, beyond the surface level, is always essentially ambiguous.

⁵ For an insightful discussion of the cognitive side of rational habituation, see Merritt (2018).

particular case. When a person does something, having both duty and another incentive for doing it, there can be no determinate fact, about that particular action, identifying the reason or incentive *from which* the person does it. Kant asks instead a different question: Is the agent the *kind of person* for whom the moral incentive is “sufficient” for the dutiful action?⁶ Would the person still act dutifully if all other incentives were opposed to duty? We must strive to be a person for whom the answer is always *yes*. But this is not a question about something that occurs in us at the time we act.

The incentive from which I could be said to act depends on my “disposition” (*Gesinnung*). This term implies that I am disposed or “minded” (*gesinnt*) to act in a certain way. Acting that way belongs, explicitly or implicitly, to my mind or consciousness (*Sinn*). This does in a way depend on my free choice because it depends on rational habituation, but I do not choose who I am *ex nihilo* every time I act. In Kant’s view there is a maxim governing my incentives: in fact, this is my “fundamental maxim.” It is my fundamental maxim that gives me “*a good or an evil heart*” (R 6:29). But that maxim is not chosen at any determinate time or chosen directly along with each of my empirical actions. My heart is evil if I am disposed to subordinate the moral incentive to nonmoral ones; my heart is good if my fundamental maxim subordinates all nonmoral incentives to the moral one (R 6:36). But my fundamental maxim is not an occurrent choice at a time. This is precisely why Kant says we cannot determine whether a human being is good or evil by directly observing particular actions (R 6:20).

The choice between a good heart and an evil heart is a “choice” in the sense that it is an act of freedom. Kant refers to it as a noumenal or intelligible choice. By this he does *not* mean that it is an occurrent choice I make in some supernatural realm of ectoplasmic intelligences. (That would be transcendent metaphysical nonsense.) He means instead precisely what he says: it is a choice that *should not be considered* in relation to time (R 6:25, 31, 39–44, 48, 49–50 n 67).⁷ It is to be thought not as an origin in time but only as an origin in reason (R 6:39) and as a single choice that applies to our use of freedom universally (R 6:25). It is not a *deed* because it *precedes* every deed (not in time but in reason) (R 6:31). Therefore, Kant’s account of the change of heart must be considered in two aspects: first, without respect to time: the change of heart or *virtus noumenon*; and, second, in its temporal manifestation: *virtus phaenomenon*.

This dual aspect of our freedom should determine our understanding of everything Kant says about acting from duty. It should help us to understand,

⁶ “Sufficiency” here could not mean *rational* sufficiency – that is, whether the moral incentive offers us a *good enough reason*. The moral incentive is *always* necessarily sufficient in that sense. The question is whether we have sufficient virtue (moral strength of character) to do what we know we have a rationally sufficient moral reason or incentive to do.

⁷ See Korsgaard (1996), pp. 181–183; and Sussman (2001), pp. 220–222; cf. also Sussman (2005), p. 173.

for instance, what Kant means when he says that the duty to act from duty is a “*narrow* and perfect one in terms of its quality, but it is wide and imperfect in terms of its degree, because of the frailty (*fragilitas*) of human nature” (MS 6:446). The final subordinate clause suggests that the duty to act from duty is best conceived in relation to the first or lowest degree of evil, which Kant calls “frailty.” Frailty is displayed when one adopts a good maxim but fails to live up to it (R 6:29). The change of heart is a life-decision pertaining to a whole life that never fully lives up to it. In §§ 4–6 of this chapter, we will try to understand better what this means in relation to the duty to act from duty. In doing this, our first task must be to address the deplorably common misunderstandings of Kant’s best-known treatment of the theme of “acting from duty” – the one found early in the *Groundwork*.

4.4 ACTING FROM DUTY IN THE *GROUNDWORK*

The early pages of the *Groundwork*, first section, is Kant’s first, most conspicuous, most often discussed treatment of the theme of acting from duty. Kant famously claims that an action would have moral worth, and its maxim moral content, only if it would be not merely in conformity with duty but also done from duty (G 4:397–399). We learn later in the *Groundwork* (G 4:407) that owing to self-opacity and self-deceptiveness, we can never be sure whether any particular action is ever done “from duty.” But in the first section, Kant wants to solicit our agreement with his claim that actions would have a special moral worth only if they were done from duty. The proposition Kant wants us to entertain involves saying that a particular action is or is not done from duty; yet Kant does not want this to depend on judgments about the agent’s disposition or character. In the *Groundwork*, therefore, it is essential to Kant’s aim that we at least have a clear criterion for saying of certain specific actions that they are done from duty and of others that they are not. We have just seen, however, that the very concept of a specific action done from duty is always problematic because the incentive from which we act is not a property of specific actions – but relates action to the agent’s disposition or character. Therefore, in the *Groundwork* Kant needs a different concept of “acting from duty” from the one used elsewhere.⁸

In the *Groundwork* Kant takes special care to *explain what he means there* by the distinction between actions in conformity with duty (*pflichtmäßig*) and actions from duty (*aus Pflicht*) – as he is using this distinction for the purposes of that discussion. In two brief pages, Kant explains the meaning of “acting from duty” by considering four examples: *first*, the honest merchant who is said to act not from duty but with a prudential or self-seeking aim; *second*, the duty to

⁸ A more detailed defense of what I say in the present section is to be found in Wood (2014), chap. 1. See also Wood (1999), pp. 26–49, Wood (2008), chap. 2. Another recent discussion of Kant’s views on motivation that agrees broadly with what I am saying is Guyer (2015).

preserve one's life; *third*, the duty to be beneficent; and, *fourth*, the indirect duty to promote one's own happiness. The second, third, and fourth examples have two variants: in the first variant, the agent acts in conformity with duty but is said not to act from duty because the agent also has an incentive of immediate inclination to act dutifully; in the second variant of these examples, that nonmoral incentive is removed (in the second and fourth examples, it is even replaced by a nonmoral incentive to act contrary to duty). Then in the second variant of each example, Kant reports that the agent still acts in conformity with duty, and he tells us that now, in contrast to the first variant of the example, the agent has acted from duty, so that the action has moral worth and the agent's maxim has moral content.

We should find it striking that in these examples we are told only: *first*, that the agent is conscious of acting in conformity with duty, from which we may infer that the agent is aware of duty as an incentive to perform this action; and, *second*, we are told what *other* incentives, if any, are available to the agent for acting dutifully. The obvious conclusion is that in this passage what he means by "acting from duty" *depends solely on the incentives that are stipulated as available to the agent*. Unfortunately, readers have tended to ignore this and have imported into their reading of Kant's examples precisely the sort of information Kant carefully excludes. They either import a non-Kantian moral psychology of motivation or they assume facts about the agents in question that would be relevant if Kant had used the phrase "from duty" in the sense he does in later works. In either case, they misunderstand what he means *in this passage* by "acting from duty" and then also misunderstand his claim that only actions done from duty have authentic moral worth.

In the last three examples, but not the first, two variants are considered. In the first example, and in the first variant of the other three, we are told that some nonmoral incentive for doing the dutiful action is also present. For the merchant in the first example, this other incentive is commercial self-interest; for the other three, it is some immediate inclination (love of life, sympathetic feeling, desire for happiness). From only this much information, Kant infers that the merchant *does not act from duty* and so his action has *no moral worth* and that the same is true in the first variant of the other three examples. Kant says he will not need a second variant in the case of the first example because it is easier to distinguish the incentive of duty from that of self-interest than to distinguish the incentive of duty from immediate inclination. This difference is due to the fact that the incentive of self-interest represents the action as something to be done not for its own sake but as a means to something else (namely, the merchant's self-interest). But Kant is supposing that both actions done from immediate inclination and actions done from duty are done for their own sake, not as a means to something else; it is because the second, third, and fourth examples involve actions done for their own sake that we need a second variant.

In the second variant of the second, third, and fourth examples, the immediate inclination is now assumed to be absent; in two of the examples (the duty to preserve one's life and the indirect duty to promote one's happiness, but not the duty of beneficence) it is even replaced by a nonmoral incentive to act contrary to duty. But we are told that in the second variant of each example, the agent still acts in conformity with duty. *From these facts alone*, Kant concludes that in the second variant, the agent acts *from duty*, while in the first variant the agent did not act *from duty*.

What Kant is doing here is explaining to us *what he means by the phrase "from duty" in this discussion*. He means, namely, that an action in conformity with duty is said to be done *from duty* only when the agent *has no incentive to do it except the incentive of duty*. Kant emphasizes this when he says that we are beneficent from duty "when no inclination stimulates us to it" (G 4:399) and again when he says of an action from duty that "every material principle has been withdrawn from it" (G 4:400). We misread the entire discussion if we import any other assumptions about the agent in these examples than Kant's stipulations about what incentives were available to the agent for acting dutifully.⁹ When Kant claims that only actions *done from duty* have true moral worth, he is claiming that there is something special – something central to *morality* – about cases where in order to perform an action in conformity with duty, the agent must be self-constrained to the action by the incentive of duty *simply because there is no other incentive available*.

As Kant presents them, his examples could not be making any point about the character or virtue of any of the agents because we are told nothing about these matters. Kant does not say, for example, that the honest merchant is a *selfish* or a *self-seeking person*, but only that his honest action is done with a "self-seeking aim" – that is, that he has a self-interested incentive for doing it and that his honest action, in fact, serves this aim. We totally miss Kant's point if we think we are to suppose that the honest merchant is not really an honest person but rather someone who would cheat customers if he did not have the incentive of business prudence for his honest dealing, or if we think that the sympathetic friend of humanity is a shallow and capricious person who would behave callously toward people in need if he were in a bad mood. Kant never says or implies any such thing about these agents; such assumptions (commonly made in discussions of this passage) disfigure the examples and guarantee that we will miss Kant's point. Still less could Kant be making any point about the "real motive" from which an agent might be acting when more than one

⁹ From the fact that each of the agents in the second variant of the last three examples still act in conformity with duty, we might justifiably conjecture that the agent is at least virtuous enough to be able to constrain himself to do this particular duty when no incentive other than duty is present. But the only mention of character in this passage is an addendum to the third example (the example of beneficence); it goes outside the scope of these examples to make the general point that character, as shown in moral self-constraint, is different from temperament (G 4:398).

incentive is present (cases some describe as “motivational overdetermination”). We are never given any example in which an agent is said to act from duty because the incentive of duty is preferred over some other incentive. Nor could Kant be making any judgments (positive or negative) about the moral value or lack of value attaching to sympathy or any other inclination – except the judgment, with which he expects his reader to concur, that dutiful actions without the presence of these *nonmoral* incentives have that special worth that is most central to *morality*.

This passage early in the *Groundwork* is one of the most commonly misunderstood passages in all of philosophy. Whether Kant’s readers would make the judgment he expects of them is seldom clear because they usually misunderstand the proposition Kant is trying to put before them. The commonest misreading takes him to be saying something negative about inclinations as incentives. It usually focuses especially on the example of beneficence from sympathy. For Kant it is no moral criticism of sympathy, or of actions done from sympathy, to say that sympathy is not the incentive most central to morality. It might, for all that, be an incentive that ought to be cultivated because it serves important moral ends; in fact, that is precisely what Kant holds (MS 6:456–458). Sympathy is one of the inclinations supportive of morality. Kant even takes it to belong to what he describes as *virtus phaenomenon*. It is profoundly sad that utterly misaimed criticisms of this passage account for so much of the literature on Kant’s moral theory.

The sole aim of Kant’s discussion at the beginning of the *Groundwork* is to make the point that dutiful actions have a special kind of moral worth, a worth central to morality, when circumstances are such that if they are to be done at all, they *have to be* done with rational self-constraint from respect for the moral law. The distinctiveness and moral importance of such cases are then used by Kant to derive a formula of that law the respect for which would figure in such cases of moral self-constraint (G 4:400–402). Kant’s discussion might serve his immediate purpose very well if only it were not so commonly and so disastrously misunderstood.

4.5 LEGALITY AND MORALITY

In Kant’s abstract and simplified usage in the first section of the *Groundwork*, it suffices for the action *not* to have been done from duty that other nonmoral incentives for the dutiful action were present to the agent. Kant’s use of the phrase “from duty” in that discussion would require us to deny that some actions are done from duty that in other contexts he might want to say are done from duty. This applies, for instance, to Kant’s distinction between the mere legality of an action (its conformity to duty) and the action’s *morality* (when the agent’s disposition places the incentive of duty ahead of other incentives). A dutiful action might be said to be done from duty in the sense that gives it *morality* if the agent’s character or disposition is such that the same

agent certainly would still have done the same dutiful action, under the same circumstances, even if all nonmoral incentives were removed. This seems a straightforward account of “acting from duty” in the context of the distinction between legality and morality. But in the context of Kant’s thesis of radical evil, such a claim is nevertheless deeply problematic in two prominent respects.

First, Kant thinks we can never know anyone’s disposition well enough to be sure that actions have morality as well as legality. In part this is because our character and incentives are empirically *self-opaque*. The source of this opacity is not metaphysical but (in Kant’s sense of the term) *anthropological*.¹⁰ Evil, as we have seen, always has its source in self-deception about who we are and therefore about the rational priority of reasons or incentives. It is part of the propensity to evil in human nature that we are disposed to conceal from ourselves why we act as we do, to misrepresent our actions, character, and disposition in a self-flattering light. For this reason Kant holds that we have a moral duty of self-examination. We are required to know ourselves well enough to carry out the inner process of self-judgment that Kant calls “conscience” (MS 6:441–442). The fulfillment of this duty is always difficult and problematic, due to the empirical difficulty of self-knowledge and our constant temptation to self-deception, but we are capable of carrying it out sufficiently well to make judgments of conscience possible. The problematic lack of self-transparency about our disposition has important consequences for the duty to act from the motive of duty, as well as for the change of heart.

The *second* difficulty, however, arises from what we have called Walpole’s thesis: “All men have their price.” In relation to our present concerns, however, we might better state Walpole’s thesis this way: “None of our actions is ever done *simply and solely* from the motive of duty.” Either the incentive of duty must be assisted by other, nonmoral incentives or at least it must not be opposed by contrary nonmoral incentives that are sufficiently great that we would succumb to the temptation to violate duty with which these nonmoral incentives present us. This is one especially strong way of putting the thesis that our propensity to radical evil is inextirpable (*unvertilgbar*). We can still say an action is done from duty, even “solely” from duty, if (as Kant does in the *Groundwork*) we assume (or stipulate) that there is no other incentive present (in favor of the action or opposed to it). But as Kant points out in the *Groundwork* itself, owing to our self-opacity, we can never know whether such an assumption holds in fact (G 4:407).

¹⁰ Our self-opacity could not possibly lie in any imagined distinction between a noumenal self in an unknowable world of things in themselves and a phenomenal self with whose actions we can be empirically acquainted. If we were completely self-transparent empirically as regards our character and motivation, then it would be absurd to claim self-opacity based on a fancied distinction between a knowable phenomenal self and an unknowable noumenal self. For a discussion of the issues surrounding Kant’s account of the difficulty and also the duty of self-knowledge in relation to conscience, see Ware (2009).

Suppose King Henry VIII wants to execute his innocent wife Anne Boleyn by falsely accusing her of adultery and treason; and suppose you are a courtier in the English court. Henry offers you money if you will falsely accuse Anne of these misdeeds. You may very well be morally strong enough to resist that temptation. But now suppose, as actually happened to several English courtiers, Henry then threatens you with death if you refuse to tell the lies he commands you to tell. Would you then still be morally strong enough to do your duty? In the second Critique, Kant uses this example, and our reaction to it, to persuade us that although we may be uncertain what we would in fact do, we would realize that we still *can* do our duty even under these extreme circumstances (KpV 5:30). (This is the paradigm case of “Ought implies Can,” as Kant understands it.) Kant thinks that an uncorrupted ten-year-old boy would spontaneously admire someone who did their duty in such a case and also wish that if he were so unfortunate as to be in the same predicament, he would also have the moral strength to do his duty, even if he had to remain unsure whether he would actually have the virtue needed to do it (KpV 5:155).

Kant’s example was not fictional. In his great *History of England*, David Hume reports that Henry Norris, Groom of the Stole (1482–1536), when put in just such a position, declared that he would sooner die a thousand deaths than calumniate an innocent person.¹¹ Norris bravely did, even in that extreme case, what “Ought implies Can” says we all *can* do. Kant insists that we would all judge that we (and Henry Norris, too) still *can* resist King Henry’s bribes and threats, no matter how high the price. But Walpole’s thesis (as Kant interprets it) would imply that if King Henry had raised the price of doing his duty far enough, then even Henry Norris would eventually have told the lie. Walpole’s thesis, as Kant reads it, asserts that none of us would *in fact* still resist temptation if the price to us were sufficiently high, even though we all still *can* resist.

As we saw in Chapter 3, Kant refrains from directly asserting Walpole’s thesis, leaving each of us to decide whether we agree with it (R 6:39). But Kant obviously thinks we ought to take Walpole’s thesis seriously for purposes of moral discipline. We should apply it chiefly in our conscientious self-examination of our own conduct, but we should not employ it to our own advantage in our attitudes toward others. We ought always to hesitate before looking down on those we see behaving worse than we do: “How many people who have lived long and guiltless lives may not be merely fortunate in having escaped so many temptations?” (MS 6:392–393).

Still another way to put Walpole’s thesis would be this: our radical propensity to evil is strong enough that no one’s disposition is ever going to be able to withstand every possible temptation to violate the moral law. That means that in every case where I do my duty, either I have nonmoral incentives for doing it or at least I am fortunately free of contrary nonmoral

¹¹ Hume, 1983, vol. 3, chap. XXX, 1, p. 118.

incentives in whose presence I would no longer do as I ought. From this we can draw some significant conclusions about the Kantian distinction between legality and morality. One is that if someone were to think that an action done *from duty* must be an action by someone for whom the thought of duty is truly the *only* incentive that plays any role in the performance of the dutiful action, then in the sense of “from duty” relevant to legality and morality, Walpole’s thesis entails that in that sense no human being ever would act from duty, and no action would ever have morality as well as legality. Consequently, if we say of any action that it is done from duty in the sense that it has morality rather than legality, we cannot mean that it would be done from duty even if all nonmoral incentives were removed. The most we might say of any of our actions is that we would still do it if all *actually present* empirical incentives were removed.

From these considerations we may further infer that any claim about a particular action that it has morality as well as legality must always be qualified in two ways: *First*, I cannot be certain that the judgments about my character or disposition on which it rests are not self-flattering inner lies. *Second*, even if I would still do my duty if these specific inclinations to do my duty were not present, Walpole’s thesis tells me that there are always going to be inclinations contrary to duty in whose presence the incentive of duty would no longer be sufficient and I would act contrary to duty. Even the *meaning* of the claim “This action of mine has morality and not mere legality” is therefore never clearly determinable because we do not know relative to which temptations our action is being judged. Further, in whatever way that meaning might be further determined, we can never know with certainty in any particular case whether the claim is true.

What do these uncertainties imply about the meaning of Kant’s claim that we have a duty to act from duty? We will now see that they require us to distinguish two aspects of this duty: it is both a narrow or perfect duty and a wide or imperfect duty. The meaning of the change of heart, central to the *Religion*, depends on these matters.

4.6 THE NARROW DUTY TO ACT FROM DUTY

We have seen that when Kant describes the duty to act from duty, he refuses to classify it simply as either a perfect or an imperfect duty. Again: “This duty to oneself is a narrow and perfect one in terms of its quality; but it is wide and imperfect in terms of its degree, because of the frailty (*fragilitas*) of human nature” (MS 6:446). I now suggest that this somewhat obscure characterization of the duty to act from duty can be greatly illuminated by its relation to the “change of heart” as presented in the *Religion*.

Kant says, “Virtue is always *in progress* and yet always starts *from the beginning*” (MS 6:409). For a human being self-afflicted with a propensity to evil, moral progress is a struggle against that evil propensity. The struggle

must *begin* with a principled decision to oppose the propensity to evil, but the struggle must be carried on over time, even over an entire future lifetime, through a laboring and striving to increase one's moral virtue. As we saw earlier, moral reform must always be a two-stage process. But the stages do not represent a temporal order; they consist in a distinction between a moral revolution – a *qualitative* transformation of our heart, which is not considered in relation to time at all – and a gradual progression in *degree* of virtue (*virtus phaenomenon*) in our life-conduct throughout our future.

The first or fundamental stage – the *beginning* from which virtue always starts – is the *change of heart*. It is grounded on a return to our fundamental conception of ourselves as self-governing rational agents, members of a community of such agents as ends in themselves, for whom the moral incentive – goodness of heart – is the ground of all rational incentives. That overcoming of self-deception about who we are transforms the human being's "attitude of mind" or "way of thinking" (*Denkungsart*). This constitutes the (perfect) *qualitative* aspect of the human being's fulfillment of the duty to act from duty. This is the adoption of what Kant calls the *maxim of holiness*: that the moral incentive is to be given absolute priority over all nonmoral incentives. In the *Religion*, Kant describes the change of heart as the acquisition of a "new heart" (R 6:51). In even more dramatically Christian terms, it is a "rebirth," the "putting off of the old man" and "putting on of the new man" (R 6:47, 73, 117, 121, 163, 202). The change of heart, Kant says, is a duty (R 6:66–67): it is a *narrow or perfect duty* because it is a "single unalterable decision" (R 6:48), without which no moral reform would be possible.

If the change of heart and way of thinking is genuine, its genuineness is attested only by the second stage, involving changes in the human being's conduct as the person's course of life unfolds in time. This is the aspect of virtue that is a matter of *degree* and always *in progress*. Because the human being is still morally imperfect (always under possible threat, according to Walpole's thesis, from opposing incentives), and since the propensity to evil is inextirpable, the holy maxim represented by the change of heart cannot be carried out all at once and can never be carried out completely. This feature of the human being's moral striving manifests the *frailty* of human nature, pertaining to the propensity to evil. There is always a gap between adopting the maxim of a good heart and actually following this maxim in the conduct of one's life. If the change of heart is genuine, it will show itself in a changed attitude of mind, which gradually brings about changes in habits, desires, inclinations, and feelings. All these gradually cease to be incentives in service to the propensity to evil that compete with the moral incentive. Instead, they come to manifest a contrary strengthening of the human being's moral character in service of the moral incentive. This is *virtus phaenomenon* as a true manifestation of *virtus noumenon*.

The acquisition of virtue happens through the strengthening of those feelings that are a direct response to reason: moral feeling, conscience, love of human beings, respect for oneself and others and for the moral law (MS 6:399–403, cf. G 4:402 n, KpV 5:71–89). It also displays itself through the strengthening of certain empirical feelings, such as sympathetic participation in the lives and situation of others, which provides not only incentives to help them but also the understanding that makes your help both effective and properly respectful of their human dignity (MS 6:456–457). Kant also considers genuine Christian love to be among the empirical feelings strengthened by the change of heart (ED 8:337–339). Of course, empirical feelings and desires in service to the evil propensity must always still be resisted; but rational feelings, and also good empirical feelings, constitute many different ways that a strong or virtuous moral character can resist them.

Kant therefore distinguishes between the two kinds of moral reform we have already identified. *Virtus noumenon* is the *qualitative* revolution from the maxim of an evil heart to that of a good heart (R 6:46–47). It is the “single and unalterable decision” to “reverse the supreme ground of all maxims” (R 6:48). Any human being, however, is always still beset with a propensity to evil that is inextirpable. This propensity shows itself in “the frailty of human nature”: performance always lags behind resolution. Our life-conduct can therefore display at best *virtus phaenomenon*: a gradual progress in degree of virtue as manifested in the empirical character, including our feelings and inclinations (R 6:47). *Virtus phaenomenon* may always be merely a semblance or false appearance of virtue, but it is also the only possible manifestation of true virtue. As the human being’s character becomes more virtuous, the human being’s conduct gradually becomes more reliably dutiful, able to withstand greater and greater temptations to transgress duty. Thereby this conduct also exhibits a strengthening of the incentive of duty, which now is increasingly able to defeat nonmoral incentives when they come into conflict with the moral law. These changes also constitute a striving toward a character for which the incentive of duty would always be sufficient to secure obedience to the moral law. The imperfect human being never reaches such a state of perfect holiness, but the progressive striving in that direction constitutes the other main aspect of the human being’s fulfillment of the duty to act from duty – the compliance with this duty as regards its *degree*. This aspect represents the duty to act from duty as a wide or imperfect duty. It is a duty to set an end: *holiness* or the *sufficiency* of the moral incentive under any and all circumstances and temptations. The end can never be attained by us, but particular strivings toward it are meritorious; their absence is not blameworthy unless it involves not setting the obligatory end at all.

True moral reform, therefore, requires the *qualitative* change of heart: a new “attitude of mind” (*Denkungsart*), making the human being into “a subject receptive to the good. But he becomes a good human being only in incessant laboring and becoming” (R 6:48). This laboring is not merely a “habituation

(*assuetudo*) – that is, uniformity in action that has become a *necessity* by repetition” – but also a *rational* and free habituation (*habitus libertatis*) that creates “an *aptitude* [*Fertigkeit*] (*habitus*) that is a facility [*Leichtigkeit*] in acting and a subjective perfection of *choice* [*Willkür*]” (R 6:407). A *habitus* or aptitude in this sense is something that makes easier what would have been more difficult without it. And an aptitude that perfects our power of choice is one that makes it easier to act not merely *legally* (in conformity with duty) but also *morally* – that is, to acquire a character such that the incentive of duty is closer to being a sufficient incentive. Because the evil propensity is inextirpable, the human being’s conduct in carrying out the resolve, and even the person’s character, always lag behind that resolve: “between maxim and deed there is still a wide gap” and even the best human being is always “upon the road of endless progress toward holiness” (R 6:46).

4.7 THE WIDE DUTY TO ACT FROM DUTY

The second aspect of moral reform – the gradual progress – puts us in a position to understand Kant’s account of the duty to act from duty “in terms of its degree” – that is, as “a wide or imperfect duty in combat with the frailty of human nature” (MS 6:446). Kant describes the duty to act from duty, regarded as a wide and imperfect duty, in the following way: it is “the duty of assessing the worth of one’s actions not by their legality alone but also by their morality (one’s disposition).” This is a duty only of wide obligation because “the law does not prescribe this inner action in the human mind.” It is not a duty to perform some inner act, either of assessing every single act in the right way or of choosing the incentive by which one acts. Instead, you habitually assess your actions by their morality, not their mere legality, and you might focus on particular acts only when they are problematic. You can’t choose the motive of this action (as if the motive were some causal lever producing it). The duty to act from duty is therefore “only the maxim of the action [– namely,]to strive with all one’s might that the thought of duty for its own sake is the sufficient incentive of every action conforming to duty” (MS 6:393).

The change of heart consists in a resolution to act, in every case where duty is at stake, not only in conformity with duty but so that the incentive of duty is a sufficient incentive for the dutiful act. This is a resolution regarding your character and disposition: your habitual susceptibility to incentives of action. It is the general resolution that the pure incentive of the moral law should have priority over all nonmoral incentives that might tempt you to violate the law. This is not a resolution that could apply directly to any individual act because no particular action could ever guarantee a decisive change in my overall disposition. Besides, Kant holds that we ought to welcome inclinations to do what duty requires. The cultivation of such inclinations is even constitutive of *virtus phaenomenon*. It is the only empirical sign that the change of heart has actually occurred in us.

It is due to the frailty of human nature that the duty to act from duty takes the form of a wide or imperfect duty. For the fact that a human being is susceptible to an inextirpable evil propensity entails that it will never fully live up to the maxim of purity adopted by the change of heart. Even if there is a change of heart, violations of duty will still occasionally occur because the incentive of duty will never be by itself sufficient for the performance of every dutiful action. For the purpose of moral discipline, we must assume that Walpole's thesis will always continue to hold of us even as we make moral progress.

The change of heart is the basic choice by which a human being decides to become morally good. But owing to the frailty of human nature, a human being who begins from evil never fully lives up to good maxims or good intentions adopted in the change of heart. The thought of the change of heart, if it occurs at a time, does not immediately bring about a good character; at most it puts in motion a process of *rational habituation* that strengthens the human being's character and gives the pure moral incentive the upper hand over other incentives. This is the point at which the duty to act from duty becomes a wide or imperfect duty and a duty involving degree.

Kant thinks we should welcome nonmoral incentives to act in conformity with duty because they make it easier to do our duty and also more likely that we will do it. But he also warns that these nonmoral incentives to do our duty are "hazardous" (KpV 5:73) because we may become dependent on them and thus give in to the degree of evil he calls "impurity" (R 6:30). This ambiguity is the same one that characterizes the concept of *virtus phaenomenon*: The *appearance* of virtue may be only a *false semblance* of virtue, but at the same time it is the only possible *manifestation* of *true* virtue.

When we have nonmoral incentives to do our duty, we must always "abstract" from them when it is a question of "the determination of duty" (TP 8:279). That is, we should prevent them from distorting our judgment about what our duty is in a given case, since these incentives may mislead us into thinking that something is our duty just because it is favored by one or more of them. Kant thinks that when people think they are acting from duty, often what is really happening is that they are flattering themselves, acting from "some covert impulse of self-love, under the pretense of the idea [of duty]" (G 4:407). But of course the genuine idea of duty is also present to them as an incentive. For Kant it is crucial to keep moral and nonmoral incentives distinct in our thinking about morality and to keep firmly in mind that I ought to be acting from duty rather than from a nonmoral incentive – whether that other incentive be business self-interest or sympathy or love of life or self-flattery masquerading as duty.

It is always a vice to deceive yourself; deceiving yourself that a nonmoral incentive is the incentive of duty is the most crucial case of this vice. But it is a basic misunderstanding of Kantian ethics to think he believes nonmoral incentives to do your duty are simply bad. The duty to act from duty, as we have seen, has two aspects: first, the change of heart, which is the resolve to give

the moral incentive priority over nonmoral incentives. But that is *not* a duty not to have nonmoral incentives to do your duty (which might be morally disastrous if you succumbed to vice); and it implies no judgment that these incentives themselves are bad. The second aspect of the duty to act from duty is the duty to strive to make the moral incentive sufficient. Other incentives to do your duty could still be present even if it were sufficient – which it never will be if Walpole’s thesis is true of you. And as long as the incentive of duty is not sufficient, you should be glad you have other incentives to do your duty. Nonmoral incentives to do your duty are good, not bad, even if they are never what is most essential to morality and even if it is a vice to become dependent on them.

What might make it the case that I am acting from duty rather than from some sympathetic or self-flattering inclination masquerading as the incentive of duty? And how might I know when I am acting from duty rather than from such an impulse or inclination? It is important to see that for Kant, there can be even in principle no clear answer to any of these questions.¹² It’s *not* that there is an answer to them here and now, not known to us but known to God. If we suppose that God knows what we don’t know about this, then that is not because he has some special, prying knowledge into the present aetiology of our particular acts but only because God knows, and can wisely judge, the whole history of our lives, including those actions that are still to us future and contingent. But to us, who are still living our lives in time, who we will be, even who we are right now, is always still at the mercy of our freedom, still up to us and therefore still not yet decided. That’s why our moral disposition is to us always still to be determined and therefore always unknowable.

Therefore, acting from duty, or giving our actions morality as well as legality, is always primarily something we aspire to do, not something we will ever actually achieve. If our actions can be said to have morality as well as legality, then that is because, first, we have undergone the change of heart and, second, because we are manifesting it in our course of life through our striving to make the incentive of duty sufficient. But whether we have done these things, or are doing them, can never be a matter of knowledge, or even justified Belief – but, as Kant says, it is always for us a matter only of *hope* (R 6:48, 69, 76).

The change of heart, and the pure disposition that results from it, is therefore never something we can claim actually to possess or on which we can base any claim to have overcome the propensity to evil or any hope to become well-pleasing to God. That is why, as we will see in the next chapter, it is appropriately symbolized as a reality distinct from us, an ideal of humanity or

¹² “What I say, I don’t feel / What I feel, I don’t show, / What I show isn’t real, / What is real, Lord, I don’t know, / no, no, no . . . / I don’t know.” – Leonard Bernstein, *Mass*, Confiteor. In Papish (2018) chap. 7, there is an interesting discussion of the change of heart as involving a two-stage process described above involving self-knowledge and the overcoming of self-deception.

the Son of God; for those Christians who accept scriptural revelation, it is symbolized as Jesus Christ.

4.8 NOT A DATABLE EVENT IN TIME

Kant was raised as a Lutheran by parents who were both Pietists. Pietism was an evangelical revivalist movement that began in the seventeenth century and continued to influence German culture through the eighteenth century and even beyond.¹³ Many Pietists were what later came to be called Pentecostals¹⁴ – that is, they believed that salvation is possible for a sinner only through a dramatic “born again” event in the sinner’s life, through which the sinner experiences the transforming hand of the Holy Spirit descending upon the sinner with the gift of divine grace. Pietists often saw this experience of divine grace descending upon them as a “born again” experience in which the Holy Spirit was felt by individual believers, assuring them of salvation through God’s grace, involving a “revolution” from an evil heart to a good heart.¹⁵

Kant was undoubtedly acquainted with Pietists who had such beliefs; the text of the *Religion* exhibits at many points a consciousness of this aspect of Pietism as something that might appeal to some of his readers. Kant’s relation to his Pietist upbringing was, however, always ambivalent at best. All explicit references to “Pietism” in Kant’s published writings express negative attitudes toward what he takes to go by that name (R 6:184 n, SF 7: 10, 55, 57).

On the surface, Kant’s doctrine of the change of heart obviously has much in common with the Pietist belief in a sudden, dramatic conversion or “rebirth.” The first is simply the term “*Herzensänderung*” itself. An *Änderung* (alteration) occurs when a single persisting subject successively takes on different states. In the case of the change of heart, it is an alteration in the human being’s fundamental maxim: first that maxim subordinates the moral incentive to nonmoral incentives, then the human being adopts the maxim of subordinating

¹³ For sources describing the history of Pietism, see Stoeffler (1973); and Shantz (2013).

¹⁴ Pentecost was a transformative event in the early church. The name Πεντηκοστή (“fiftieth”), refers to the fiftieth day after Passover. On the Hebrew holy day of *Shavuot*, the Holy Spirit is said to have descended upon the early Christian Church, making them God’s people.

¹⁵ The term “revolution” may be significant. There was a distinctly radical social spirit to some of Pietist religion, which shows itself in Kant’s emphasis on a religious community. This emerges in part three of the *Religion* and will be discussed below in Chapter 7. “And all that believed were together, and had all things common; And sold their possessions and goods, and parted them to all men, as every man had need” (Acts 2:44–45). This is the obvious source of the slogan coined in Louis Blanc’s 1840 pamphlet *Plus de Girondins*: “From each according to his abilities, to each according to his needs.” When Karl Marx quoted this slogan in *Critique of the Gotha Program*, his audience of German socialists would have known he was quoting it from Louis Blanc. It is a sign of deplorable ignorance to attribute the slogan to Marx himself – all the more so because Marx is *not* offering it as a distributive proposal for the working-class movement. He explicitly denies that it is any sort of principle of *right*. For Marx it refers to a future society that will have “transcended the narrow horizon of bourgeois right in its entirety.”

all other incentives to the moral one. Kant also describes the change as if it were a sudden and decisive alteration: “a single and unalterable decision” (R 6:48); as we have seen, he contrasts the change of heart to the gradual process of moral reform that will follow upon it – even over the rest of the human being’s earthly life and beyond, if we can suppose the human being will continue to make moral progress in a future state (KpV 5:122–124, R 6:67–71). Since the change of heart is supposed to occur in the life of a person who has begun from evil, it is also natural to think that it takes place at some time in the person’s life. A Pietist Christian who believes salvation comes to the sinner through a sudden “born again” experience might easily take Kant to be expressing agreement with that belief.

There are, however, several features of the usual Pietist conception of this conversion from which Kant evidently wishes to distance himself. *First*, and most basically, the change of heart is not a decision to be considered in relation to time. Further: “Virtue is always *in progress* and yet always starts *from the beginning*” (MS 6:409). The beginning – the change of heart – must be constantly renewed; better yet, it must be a constant presence throughout the entire future life-conduct of the moral agent. *Second*, Kant denies, on basic critical principles, that there can be any empirical knowledge of God or divine operations and effects. He does not hesitate to apply this doctrine to alleged operations of divine grace on the human being: “The persuasion that we can distinguish the effects of grace from those of nature (virtue), or even produce these effects in ourselves, is *enthusiasm* (*Schwärmerei*); for nowhere in experience can we recognize a supersensible object, even less exert influence on it to bring it down to us” (R 6:174).

I don’t think Kant means to deny that the *thought* of the change of heart – the single resolute decision – *could* take place at a certain datable time in a person’s life (or at many such times). Kant would deny, however – supposing I am one of those in whom this occurs – that I, or any other human being, could ever know, or even have good grounds to believe (or Believe), that the change of heart occurred at just this time or in just this way. The Belief that it had occurred cannot, therefore, be part of rational religion or necessary to a saving faith. The only possible evidence that I have undergone a change of heart is to be found in my future conduct – in a life of gradual moral improvement over the rest of my days (R 6:48, 67, 76). For this reason the propositional attitude Kant most often says I must take toward my own change of heart is never *knowledge* or even *Belief* but only *hope* and never *certainty* (R 6:76, cf. R 6:48).

Again, as we will see in Chapter 5, the religious symbol for the pure disposition resulting from the change of heart is an ideal person or the Son of God, distinct from myself, on whose reality I must base my hope for salvation. Perhaps Kant’s Pietist acquaintances might say that the hand of grace touching them was the presence to them of Jesus Christ. As Kant sees the matter, that would be a symbolic way a Christian might describe it. But Kant would deny that they could ever claim to know, or justifiably believe, based on experience,

that such a change in their lives had actually occurred. They might hope it had, but then they would have to support this hope with a gradual moral progress in the future course of their conduct.

My guess is that Kant was familiar with people who, according to their own testimony, had just undergone such a conversion experience: they claimed they had repented of their sins, felt the hand of the Holy Spirit on them, resolved to lead a new life in Christ, and were overcome with joy and gratitude to God for their certain salvation. On Sunday they had felt the hand of grace on them, but then on Monday morning they went back to their old corrupt ways, with the one difference that now they had been saved and could calmly go about their same depraved business with no further fear of Hell. Clearly, therefore, we should be skeptical of such pious declarations. The attitude Kant recommends we take toward such people, however, is sooner charitable and permissive than harsh and judgmental. We should “throw a veil of philanthropy over the faults of others, not merely by softening our judgments but also by keeping these judgments to ourselves; for examples of respect that we give others can arouse their striving to deserve it” (MS 6:466). Even the deceptive (and self-deceptive) pretense of virtue should be tolerated in others, though never in ourselves: “Even the illusion of good in others must have worth for us, for out of this play with pretenses, which acquires a respect without earning it, something serious might finally develop. It is only the illusion of good in ourselves that must be wiped out without exemption, and the veil by which self-love conceals our moral defects must be torn away” (Anth 7:153).

To understand ourselves as most fundamentally free, self-governing rational agents makes a difference in how we think about people’s characters and actions. Freedom is fundamentally a matter of taking responsibility for yourself as a member of the realm of ends and treating both yourself and others as subjects and not only as objects. Freedom may be incomprehensible, but we turn it into a mere object if we are distracted by the metaphysics of the unknowable. Self-knowledge is never merely a matter of objective properties or causal interactions with things. It is a closed book to our shallow, arrogant causal natural and social sciences. Self-knowledge is always a matter of existential anxiety and self-honesty about my fundamental identity as a free rational agent, a member of the realm of ends. Empirical knowledge provides evidence that we give the moral incentive priority over competing incentives. But Kant insists that we can never be quite sure from what incentive we are acting (G 4:407). This is not so much a matter of our ignorance of fixed, objective facts as it is a relation to our indeterminate future because who we will be is always up to us through our living of our still unfinished lives. Our hope that we have undergone a decisive change of heart is not directed to something abstract and metaphysical, taking place in a different world. It is about something immediately present to us, but because it is partly future, it is also something of which we can at no time in our lives ever claim absolute

possession or certainty.¹⁶ It is an uncertainty about who I am – about my own rational agency, who I have made myself to be, and who I will become, all through my own freedom. This is why the change of heart must be something for which we must hope but of which we can never claim full or certain possession. The uncertainty is not a matter of objective ignorance and knowledge but fundamentally a struggle of inner honesty with bad faith.¹⁷ It would display not a cognitive but a *moral* deficiency to think the uncertainty about people's incentives could perhaps someday be removed by progress in brain science. That would be fundamentally the same moral defect as the one displayed by Kant's Pietist contemporaries when they deluded themselves that a religious experience could provide people with certainty that they had been transformed by divine grace.

The change of heart could not be a datable event because it is about *who we are* and therefore about the always still unfinished course of our lives as free beings. That – and not any transcendent metaphysical nonsense about a timeless noumenal world – is why Kant says it is not to be considered in relation to time. Kant's account does, however, suggest a certain stereotypical picture of what moral conversion and striving would involve. It involves a decision (either at a time or renewed repeatedly over time) – a resolve characterizing the future course of a life. And then as a sign of this decision, it involves a life of “moral progress,” which Kant seems to think of as the gradual improvement in conduct throughout the rest of a person's life. This seems to imply an acquisition and strengthening of virtues, also the combatting and laying aside of vices, but a struggle always beset with temptations and imperfection.

I don't think we would need to understand the life of moral progress in this stereotypical way in order to capture what is essential to Kant's account. What matters is that the person's conduct somehow exhibits a sustained effort to be a good human being despite the awareness of moral imperfection and persistent temptation. As we follow Kant's depiction of salvation in the *Religion*, it should help us to think about its main elements if we do not associate them with a narrative of “moral progress” in which you must be more consistently dutiful in your conduct at age forty than at age thirty and more dutiful at age sixty than at age fifty. The narrative history of each person's moral life is different from that of any other, and it will often depend on the contingencies of moral opportunity and moral temptation more than it involves the person's specific reaction to each of them. A “life of moral progress” might consist less in changes in a person's conduct than it does in the steady resolve to preserve and exercise the virtues the person has and to resist the vices of which the person

¹⁶ In his play *Murder in the Cathedral*, T. S. Eliot dramatizes a situation not unlike that of Henry Norris. He presents Thomas Becket preparing for martyrdom, still uncertain about the state of his own soul; about his identity as a free agent: “The last temptation is the greatest treason; To do the right deed for the wrong reason” (Eliot, 1963, p. 44).

¹⁷ Again, exploring the importance of these themes in Kant is the merit of Papish (2018).

remains aware. Just as the change of heart need not be thought of in strict Pietist fashion as a datable “conversion-event,” so “moral progress” might show itself in a wide variety of ways throughout the life of a person in whom the change of heart is effective. Looking at moral progress in this more flexible way would allow us to see how Kant’s account of moral progress could be relevant to more of our lives.

The Son of God

Part two of the *Religion* focuses on the sinner's difficulty in sustaining the rational hope of having undergone the change of heart and also the hope that even this moral conversion will make the sinner well-pleasing to God. From the standpoint of Kant's second "experiment," however, part two deals with a matter central to revealed Christianity: the hope for redemption through the saving work of Jesus Christ. This Christian hope belongs to the outer circle of revealed faith, which is the territory of the biblical theologian, not to the inner circle of rational faith, the domain of the philosopher. But Kant's task in the second experiment is to show how this revealed Christian doctrine leads back to the inner circle of pure rational religion.

Part two of the *Religion* begins with an introductory section in which Kant again emphasizes that the enemy against which we struggle is not natural inclination but a propensity of the human will (R 6:57). Then section one focuses on the "personification" of the good principle and its "rightful claim" over the human being. Throughout part two, the struggle of good against evil is symbolized as a legal process in which the human being is hoping to be vindicated or "justified." In section two, the struggle of good with evil is symbolically represented as a legal conflict and also as a physical combat (R 6:81). The ideal of humanity or the Son of God represents symbolically the good disposition that is the outcome of the change of heart. Accordingly, religious symbolism itself will also be a central theme in the present chapter.

5.1 THE MORAL IDEAL

It is plain enough that when he speaks of "the ideal of humanity," or "the personification of the good principle," Kant is referring in some way to "the Holy One of the Gospels" (G 4:408): to Jesus Christ as revealed in the New Testament. Kant's account of the ideal makes numerous obvious references to the person and life of Jesus – especially to his being subject to temptation, which

is nevertheless resisted, and his willingness to sacrifice himself for the sake of helping others (R 6:63–64). It is striking, however, that there are precisely zero occurrences of the name “Jesus” in the entire text of the *Religion*; and the title “Christ” (German: *Christus*) is used only twice, in both cases merely reporting scriptural statements. Neither the name nor the title belong to the inner circle of pure rational religion.

It is therefore worse than gratuitous to speak, as many writers do, of Kant’s “Christology.”¹ I submit that Kant would reserve the term “Christology” for those parts of *revealed* Christian theology dealing with the person of Jesus Christ and his redeeming work that belong only to the outer circle. Christology, in short, is the domain of the biblical theologian, not of the philosopher (SF 7:23–24, 36–60). Kant’s hypothesis in the *Religion*’s second experiment is, *first*, that these revealed doctrines belonging to the outer circle can be authoritatively interpreted by the biblical theologian so that they are at least consistent with the inner circle of rational religion and, *second*, that some doctrines of revealed religion can lead back to rational religion. The *ideal of humanity* (or “Son of God”) is the result of this experiment: this ideal belongs to that in the inner circle of rational religion to which the revealed doctrines pertaining to Jesus Christ lead back.

For Kant “ideal” is a technical term. It refers to an *idea* – a pure concept of reason, to which no experience can ever be adequate – when the idea is “personified,” i.e., represented as instantiated in a determinate individual (KrV A567–568/B595–596). The *ens realissimum* is an ideal of theoretical reason in the determination of things in general. In part two of the *Religion*, the ideal of humanity is the concept of an individual human being instantiating pure moral perfection or holiness of will. In less than two pages, Kant offers eight characterizations:

- (1) “The ideal of humanity well-pleasing to God” (R6:61)
- (2) “The personified idea of the good principle” (R 6:60)
- (3) “Humanity in its full moral perfection” (R 6:60)
- (4) “The human being well-pleasing to God” (R 6:60, 61)
- (5) “The prototype of moral disposition in its entire purity” (R 6:60, 61)
- (6) “The ideal of holiness” (R 6:61)

¹ See, for example, Sala, 2004; Pasternack, 2014, pp. 133–162; and Palmquist, 2015, pp. 28, 148. Particularly wide of the mark are the suggestions by Moors, first, that Kant specifically *criticizes* many facets of Christology and then, second, that he offers us a Christology of his own that is merely a “jesuology” (Moors, 2017, pp. 551–552). It is a fundamental misunderstanding of Kant’s entire project in the *Religion* to think that he is *criticizing* anything in revealed Christianity simply because he does not include it in the sphere of rational religion. Firestone and Jacobs (2008, pp. 155–170) speculate that the ideal of humanity is not Christ but a “Platonic prototype proceeding from God’s own being,” which plays a role in medieval (especially Scotistic) and early modern Christology. These speculations have textual support only if we suppose Kant has these historical allusions specifically in mind, which seems to me highly doubtful because it would involve transcendent metaphysical commitments inconsistent with the critical philosophy.

- (7) “The God-like human being” (R 6:61)
- (8) “The Son of God” (R 6:61)

These formulations serve various purposes in Kant’s exposition. (1), (2), and (6) emphasize that the ideal is thought of as a *human individual*. (4), (5), and (6) indicate that it serves as a standard of perfection for our moral striving – even though Kant also emphasizes that since the ideal of humanity needs no effort or change to attain to holiness, the ideal *cannot* be thought by us as “an example to be emulated” (R 6:64). (7) and (8) connect the ideal to its specifically Christian sources. (3) further connects the ideal of humanity with Kant’s doctrine in the third Critique that the moral perfection of humanity is to be regarded as the ultimate end of creation and also the final end that connects all teleology, natural as well as moral (KU 5:429–436).

Decisive formulas regarding Kant’s aims in part two are (1) and (4). For they indicate the end set for the moral individual by the task of effecting the change of heart: to become “well-pleasing to God” (*Gott wohlgefällig*, R 6:60–64.) The reference is biblical: “And a voice came from heaven saying, Thou art my beloved son, in whom I am well-pleased” (Mk 1:11, cf. Gn 4:4, 1 Sam 18:26, Ps 18:19, Is 53:10, Mt 3:17, Lk 3:22, 2 Pt 1:17). I will use the terms “ideal of humanity” or more often “the Son of God.” The aim of this chapter will be to understand the role of this ideal in the rational human hope for salvation.

5.2 SYMBOLS AND ANALOGY

The struggle against evil is presented by Kant as a conflict, whether legal or physical, between the evil principle, in the form of the human being’s radical propensity to evil, and the good principle, found in the change of heart. Part two is filled with images of strife and combat: at the opening of section one, of violent physical or even military struggle (R 6:57–60) and in section two a conflict of right, in which the evil principle asserts a legal claim over the human being and the good principle acts as the human being’s advocate and defender (R 6:78–82).² The protagonist of part two is the moral ideal – “the personified idea of the good principle” (R 6:60). The personification of the moral ideal, as well as the imagery of struggle or legal process, suggests that Kant’s thought in this part of the *Religion* is carefully and self-consciously devised to be *symbolic* in the sense introduced in Chapter 1. So let us first recapitulate and expand that account.

Kant holds that all cognition requires both understanding (concepts) and sensibility: “Thoughts without content are empty, intuitions without concepts are blind” (KrV B 75). It is less often appreciated that for the application of a priori concepts in the practical sphere, Kant equally requires in addition to an intellectual element also an intuitive one, somehow related to sensibility, despite

² This theme is insightfully discussed in Chignell (2010a). Helpful also is Palmquist (2015), chap. 6.

the fact that a pure concept, especially an idea (or an ideal), cannot be represented sensibly. The presentation (*Darstellung*) of an a priori concept Kant calls a “hypotyposis” (KU 5:351). Only in this way can our concept of God come to *signify* anything to us (Anth 7:191). Hypotyposis can occur in two ways: in the case of categories that can be applied cognitively to objects of experience, it takes the form of *schemata* or *direct* presentations, which in the first Critique, Kant expounds in chapter 1 of the Analytic of Principles (KrV A137–147/B176–187). In the case of an idea, such as the idea of God, schematism is not possible. There hypotyposis must take the form of an *indirect* or *symbolic* presentation. “Symbols are merely means that understanding uses to provide the concept with significance through the presentation of an object for it. But they are only indirect means, owing to an *analogy* with certain intuitions to which the concept can be applied” (Anth 7:191).

Symbols work by way of *analogy*. Kant emphasizes, however, that analogy (in this sense) is misunderstood if it is taken as requiring a *resemblance* between the intuited object that serves as a symbol and the object of the concept to which the symbol gives significance. “This type of cognition *according to analogy* . . . does not signify, as the word is usually taken, an imperfect similarity between two things, but rather a perfect similarity between two relations in wholly dissimilar things” (P 4:357, cf. KU 5:464, VpR 28:1023). More specifically, the analogy involved in religious symbolism involves sameness in *our way of thinking* (*Denkungsart*) about things that are dissimilar or at most only imperfectly similar. Where analogues are similar in certain respects, the similarity may even play a role in the sameness of our way of thinking. But in Kant’s sense of analogy, the similarity between analogues, where it exists, is not relevant to the analogy except insofar as it accounts for the sameness in our way of thinking.

Kant illustrates the analogy involved in symbolism with this example: a fully just state or monarchy may be symbolized by an organism, whereas a despotic state is symbolized by a hand mill. What is the same here, Kant says, is “the rule of reflection” employed in thinking of the objects that symbolize the two kinds of state (KU 5:352). The similarities here matter only to the extent that they support the thought that the just state is governed by laws internal to the general will of the entire people, whereas the despotic state is governed by a single absolute will that is external to what it governs. When we reflect on this difference, we employ the same rule used in thinking of an organism as governed by a harmonious life common to all its parts, whereas we think of the operations of a hand mill as governed by the external force that turns the crank.

Kant also employs analogy in treating of our duties in regard to nonrational animals. Kant regards animals as analogues (though not symbols) of human persons. A duty toward (*gegen*) a being involves the subjection of our rational will in some respect to the rational will of that being insofar as we are empirically acquainted with it (MS 6:417). We therefore can have no duties *toward* God, natural beauty, or nonrational animals. Not toward God because

although he has a rational will, we cannot be empirically acquainted with it. Not toward natural beauty or nonrational animals because they have no rational wills but (in the case of animals) at most desires, preferences, pleasures, and pains and animal choice (*arbitrium brutum*) (MS 6:442–443). The only rational will that can impose these duties on us, Kant argues, is our own; so duties in regard to animals are really duties to ourselves. However, we can (and should) think about animals, in certain respects, the same way we think about rational beings, and that makes them analogues of the other rational beings toward which we have duties. We should think about a nonrational being who cares, desires, and suffers in the same way we think about human beings who care, desire and suffer, and therefore treat the suffering, desiring, or caring animal similarly, recognizing duties in regard to the animal that arise from this same way of thinking.

Further, the faithful service of a dog or horse, Kant says, is analogous to the devotion of a human friend or servant; likewise, the care we see animals bestowing on their young, or others of their kind or even on humans, is analogous to human love and caring (VE Collins 27:459, VE Vigilantius 27:710). This is not because their devotion or caring are *like* human devotion and caring (though they may be similar in some respects). It is rather because we ought, in certain relevant respects, to think about animal devotion and caring in the same way we think about human devotion and caring. Thus we should not cast the aged dog or horse aside like a worn-out tool but instead should repay the animal's devotion and service with gratitude.

Kant also claims that beauty is a *symbol* of morality. This is not because a flower is similar to an act of honesty or generosity but rather because the immediacy, disinterestedness, freedom (from a determinate concept), and universality of pleasure in the beautiful involves the same way of thinking about it that we think about the claims of morality on us and those dispositions and actions that respond to these claims (KU 5:353–354). The sameness in the way of thinking has the consequence that the appreciation of natural beauty is “the mark of a good soul” (KU 5:299). It also means that the appreciation of natural beauty has a moral significance for us: we ascribe to beautiful objects analogues of moral qualities – forests are majestic and magnificent; fields smiling and joyful; colors tender, modest, or innocent – because they arouse the same states of mind in us (KU 5:354).

Symbolism is needed in the exhibition or presentation (*Darstellung*) of all pure rational ideas. As we have seen earlier, the idea of God, as it is given to us a priori in metaphysics (KrV A571–583/B599–611) contains only pure categories; this gives us only what Kant calls a “transcendental theology” or an “ontotheology” (VpR 28:999–1001, 1012–1020). These concepts of God's attributes provide only what Kant (in his technical use of this term) calls a “deistic” concept of God (VpR 28:1001). As we saw in Chapter 1, a *deistic* concept of God provides no cognition of the real possibility of God even from a practical point of view and is therefore useless for religion (KU 5:353, cf.

R 6:111, 136). If the idea of God is to be applied in religion, it must be furnished with some kind of intuitive content or presentation.

5.3 SYMBOLS AND RELIGION

Obviously there are many possible respects in which we can think about two things “in the same way.” To understand any analogy, and therefore any symbolism, we must identify the respects in which an analogue is thought of *in the same way* as the object on which the analogy is based. The “sameness” here could not be mere abstract identity; in fact, this “sameness” is not even a symmetric relation. If the function of the symbol is to enhance our practical cognition of the object symbolized, the symbol must add to our thought of what is symbolized those thoughts and feelings that can accomplish this, and the sameness in way of thinking must achieve this enhancement of meaning and emotion. Symbolism involves the faculties of imagination and reflection that make possible aesthetic experience and arouse the feelings and emotions that connect our thinking with rational action.

The cases of symbolism described above illustrate this point. They express our attitudes toward just states and despotic states; they express feelings of caring toward animals; they enrich moral feelings through our experience of beauty. In this way symbolism always involves a double focus: on the *object* symbolized and on the *moral (or religious) meaning* the object acquires for us through the symbol. In the case of God, analogy has the function of providing the pure rational concept of an *ens realissimum* with practical significance. This is why thinking of the Deity only through the “transcendental” or “deistic” concept (devoid of the hypotyposis of symbolism) could have no practical significance and therefore no religious use.

As we have seen, the alternative to both deistic and theistic thinking about God is *anthropomorphism*. It represents God literally, not symbolically, as a very powerful being with a will like the human will: a being who governs nature arbitrarily, as a human ruler might govern a despotism. Since that concept of God could not be drawn from the rational idea of supreme perfection, it would involve a morally corrupt way of thinking: a slavish attitude of submission to the arbitrary will of a cosmic tyrant (see KpV 5:131, 136, R 6:65 n, 141, VpR 28:1002, 1045–1046, 1110, 1117). True religion, however, always depends on a concept of God that is free of empirical human imperfections, yet also has emotional significance for us. It is symbolism alone, therefore, that makes the theoretical idea of a supremely real being practically meaningful to us and thus makes true religion possible.

Closely related to anthropomorphism is *superstition*: “representing reason as if it were not subject to the rules of nature on which the understanding grounds it by means of its essential law” (KU 5:294, cf. KpV 5:136, R 6:53, 83, 101, 118, 174). Superstition is not merely false beliefs regarding the causes of events (for instance, belief in supernatural causes, magic spells, or witchcraft). Superstition

involves something deeper: the transgression of those very laws of understanding that ought to govern our faculties. Treating statutory religious rituals as substitutes for good life-conduct or making belief in witchcraft or miracles a condition of salvation is superstition.³ It counts as superstition only because such beliefs transgress the essential norms of reason and therefore violate a moral duty to oneself. Among these norms are those requiring us to distinguish ideas of reason from the empirical concepts fit to serve not as predicates for theoretical cognition but exclusively as symbols. Representing what cannot be cognized as if it were cognition is essentially a form of self-deception – the “inner lie” (MS 6:430).

Symbolic thought always raises the question how the symbol is to be interpreted. In presenting (*darstellen*) the ideas that belong to morality and constitute the essence of all true religion, “it is *enlightenment* to distinguish the symbolic from the intellectual (hence public worship from religion), the temporarily useful and necessary *shell* from the thing itself. For otherwise an *ideal* (of pure reason) is mistaken for an *idol*, and the final end is missed” (Anth 7:192). This is enlightenment because it involves thinking for oneself, taking responsibility for one’s own thoughts and actions, not subjecting oneself to the guardianship of others – to priests, ecclesiastical traditions, or the letter of ancient writings (WA 8:35). Facing up to the responsibility for thinking for oneself in matters of religion means acknowledging the difference between the symbol and what it symbolizes and also realizing that we can have religious cognition of the divine only through symbols. To mistake the symbol for the reality symbolized is to take our subjective representation of something for the object itself. Kant’s name for this mental malady is “delusion” (*Wahn*) (Anth 7:215, 275). In part four of the *Religion*, Kant deals at length with the phenomenon of religious delusion and how it stands opposed to the moral principle at the center of all true religion (R 6:170–175).

Enlightened believers are capable of detaching themselves reflectively from the letter of religious doctrines and practices and seeking after their true meaning, instead of accepting them in a slavish spirit of dogmatic verbal repetition and subjection to authority. Thus Kant insists,

³ Kant defines “miracles” as “events in the world the causes and *effects* of which are absolutely unknown to us and must remain so” (R 6:86). Kant does not deny the occurrence of miracles because it is not contrary to our concept of nature that we have limited cognition of the laws that govern natural occurrences. Therefore, we cannot preclude the possibility that God might allow nature to deviate from its customary laws in ways we cannot understand and thus it would not be superstition to admit the possibility of miracles. But for the same reason, we could not have evidence for miracles in particular cases. Kant does regard belief in “demonic” miracles (performed by finite spirits, whether “angelic” or “satanic”), as “irreconcilable with the employment of reason” and therefore presumably superstitious (R 6:86–87). He also condemns, as a blamable form of “moral unbelief,” the demand for miracles as a condition of our recognition of the authority of moral or religious precepts (R 6:84).

When it is a question of what teachers have really thought in their holy writings, one must not interpret them *symbolically* but rather *literally*; for to twist their words would be dishonest. But when it is a question not merely of the *truthfulness* of the teacher but the *truth* of the teaching, then one can and should interpret this teaching as a merely symbolic kind of representation, in which established formalities and customs accompany those practical ideas. For otherwise the intellectual sense, which constitutes the final end, would be lost (Anth 7:192).

The recognition that religious thought is symbolic is for Kant our only real protection against both idolatry and anthropomorphism – i.e., the corruption of thought about God through likening his perfections to human imperfections (KpV 5: 131, 136, KU 5:353, R 6:141–142, VpR 28:1046–1047). Perhaps religious believers who do not think reflectively of their beliefs as symbolic are nevertheless behaving permissibly (innocently), as long as this is merely a lack of reflective awareness of the actual role of symbols in their belief. It becomes something worse, however, if they deliberately substitute belief in the merely literal version of doctrines for a belief in which the appropriate symbolism plays an essential role. At that point Kant will say their belief has become delusion and superstition – the traditional religious term for which is “idolatry.” As we’ve seen Kant thinks there is a strong tendency in all popular religion toward anthropomorphism, superstition, and idolatry (see R 6:197–199, KU 5: 274–275, 459–460, VpR 28:1122–1126). The enlightened acknowledgment that religious thinking is essentially symbolic can sometimes serve to protect true religion from the danger of idolatry.⁴

Moral duty is not inherently symbolic, even if it lends itself to symbols – for example, beauty is a symbol of morality. Regarding our duties as divine commands, however, is already symbolic because to ascribe volition to the *ens realissimum* is already to employ analogy and hence to represent our duties in terms of religious symbolism. Kant’s position is that it is not a moral duty to represent our moral duties in this way. Therefore, religion is morally optional. But we will see in Chapter 7 that Kant thinks the human struggle against evil is unlikely to make progress if individuals, one by one, struggle within themselves against the evil propensity. It can be won only if they unite in an ethical community, and the human species collectively has a duty *sui generis* to itself to form such a community (R 6:97). Kant holds, moreover, that the real-life exemplars of ethical communities in history are religious – they are *churches*. We will see that for Kant these communities may be needed historically as vehicles for pure rational religion, but the historical task of humanity is to

⁴ As we observed in Chapter 1, the modern theologian who has most stressed the importance of religious symbolism is Paul Tillich. One important point on which Kant anticipates Tillich is in his recognition that the essence of religious idolatry consists in being unable to distinguish the symbol from what it symbolizes (R 6:169 n, cf. Anth 7:192). Kant’s chief concern is with the practical forms taken by delusion and idolatry, when the devotee hopes to win divine favor based on the worshipper’s self-degrading sycophancy (R 6:185, 198–199).

free pure religious faith from the vehicle, as a germ or kernel is freed from its husk. To the question whether Kant holds that morality requires symbolic religious representation, therefore, there can be no simple yes or no.

In Chapter 6 we will look at some ways in which Christian symbols involved in revealed religion can be interpreted through the doctrines of pure rational religion – since this is what Kant intends in his second experiment, as leading revealed religion back to the religion of pure reason. But rational religion itself is also symbolic. It symbolizes our moral life – our propensity to evil, our efforts to combat it, the Beliefs and hopes involved in those efforts.

What is the relation between religious symbols and the rational moral truths that receive symbolic form in religion? John Hare speaks of Kant as offering a “translation” of Christian doctrines (Hare, 1996, pp. 34, 40, 52–55, 62–66, 190–196). Unless meant only as loose talk, this term betrays a serious misunderstanding of the symbolic in religion. A *translation* assumes something like two distinct languages, each fully intelligible in itself. Translation is the presentation, as far as possible, in a target language of something equivalent in meaning that is said in the original language. Symbolism, however, often takes place within one and the same language; its point is to express something in that same language that cannot be fully or properly expressed without analogy and symbolism. Neither employing a symbol nor interpreting a symbol involves *translating* symbolic language into another language. An interpretation of a symbol is an attempt to make the symbol more intelligible to reason. It is *religiously* important to realize that religious thought is symbolic and that its symbols require interpretation.

Hare criticizes Kant’s “translations” as “inadequate” or “unhelpful.” This in effect assumes that religious symbols have a transparent meaning in their literal form and need no interpretation. In that case, any attempt to “translate” them into “rational” or “philosophical” doctrines would be superfluous. The *interpretation* of symbols is often not merely “helpful” but indispensable to the responsible appropriation of their meaning, while accepting them literally evades this task. It is a task imposed on any religious person who also seeks *enlightenment* – to *think for oneself* rather than submitting slavishly to some ecclesiastical authority.

Someone might ask, *What could it mean to have faith in a symbol?* The question already misunderstands the nature of religious faith. It assumes that faith is *faith in the symbol* rather than in what it symbolizes. Kant’s position is that faith is in God, and symbolism is the only way for human beings to think *religiously* about God. All rational thoughts about God as being with understanding and will are symbolic (KU 5:383). If we removed symbolism from our thoughts about God, they would have to be either “deistic” thoughts, without religious meaning or the “anthropomorphic” thoughts of a superstitious or idolatrous pseudoreligion.

Consider in this light the following statement by Hare: “Kant’s account of knowledge requires that we can claim to know only what we could in principle experience with the senses. He is certainly at liberty to confine the word

‘knowledge’ in this way. But it will not follow that the Christian cannot responsibly claim to have experience of God or to be united with Christ” (Hare, 1996, p. 274). Hare is missing the essential point: Kant’s placing of strictures on “knowledge” or “cognition” (*Erkenntnis*) by no means represents merely a technical terminological convention. It expresses an important point of epistemology that has not only implications for the way we think about our cognition of sensible objects in nature but even moral implications for the way we can think about incognizable supersensible objects such as God. Kant’s term for the claim to experience supersensible realities cognitively is “enthusiasm” (*Schwärmerei*) (R 6:53, 174).⁵ It is both a moral and an intellectual vice.

When Christians claim that they “experience” God or unity with Christ, Kant thinks these claims can be intelligible or even correct, as long as the experience is correctly understood. It cannot be an empirical cognition. Its content must be essentially *symbolic*.⁶ Such claims and such experiences employ concepts drawn from sensibility, but what is thought using them depends on a sameness in the way of thinking, not a literal sameness in what we are thinking about. Symbolic thought implicitly takes upon itself the burden of self-understanding, acknowledging that there is something being thought that cannot be understood directly and is open to interpretation. The contrary pretense, at least if adopted deliberately and stubbornly, is religious delusion, superstition, perhaps even idolatry.

5.4 THE SON OF GOD AS A SYMBOL

Let’s now summarize what has been said by formulating a bit more explicitly the concept of a *religious symbol*:

S is a *religious symbol* of R = df. (1) R is an *idea of reason*, or a reality in our moral life

- having religious significance,
- (2) S is an intuitive representation which serves to *present* (*darstellen*) to us intuitively the idea R or its object, and
- (3) S presents R by means of an *analogy*: that is, by means of the fact that we think religiously about R in the same way we think about S.

Clauses (1) and (3) need some commentary. Clause (1) refers first to ideas of reason, which cannot be represented without hypotypothesis; to have sense and

⁵ For a good recent discussion of this important Kantian concept, see Zuckert (2010). Kant occasionally uses the term “Enthusiasmus,” whose meaning, however, is quite different. “Enthusiasmus” refers to strong emotions or affects that can sometimes produce a good state of mind and good conduct (KU 5: 272, MS 6:409, SF 7:85). This kind of enthusiasm does not necessarily involve either a cognitive error or a moral failing.

⁶ Kant appears at least to countenance, if not quite to commit himself to agree with, what he calls “a pure mysticism in religion,” as represented in a letter from which he quotes by C. A. Wilmans (SF 7:69–75).

significance for us, they must be represented symbolically. The principal idea of reason involved in religious symbolism is, of course, that of God as a being commanding our moral duties. Some empirical features of our moral life, such as our resolve to combat the radical propensity to evil, can also receive religious symbolism through implicit reference to God as commanding our moral duties.

Regarding clause (3): first, it entails clause (2) if we suppose that analogy always involves intuitive presentation; but it adds to clause (2) the thought that the hypotyposis is by way of analogy – and therefore must be a symbol rather than a schema. Second, as was said earlier, the “sameness in way of thinking” involved in analogy is an asymmetric relation: it presents what is symbolized in a way that enhances the practical or emotional significance for us of the reality symbolized. The word “religiously” in clause (3) is meant to indicate that the enhanced moral significance relates to the symbolic representation of our duties *religiously*, as divine commands.

So understood, Kant’s use of the ideal of humanity or the Son of God offered us in part two of the *Religion* involves religious symbolism in three ways or on three levels:

First, the basic moral reality of our lives that is to receive symbolic presentation is the *change of heart* and the improved future course of life to which the change of heart leads – and which is the only possible attestation of its reality. Both are something we can never be certain has occurred (or will ever occur) in our own lives. But we may Believe in their occurrence or at least hope for it. This resolve and this moral progress, where they exist, are not symbols but literal parts of the person’s moral life. However, in religion the moral striving for which we hope is represented symbolically as our striving to become well-pleasing to God. Kant also presents the change of heart symbolically in biblical terms as a “putting off of the old man and a putting on of the new” (R 6:48, 73; cf. Col 3:9–10, Rom 6:6, Eph 4:20).

Second, the Son of God is a religious symbol for the change of heart or, more specifically, for the good disposition resulting from it. This disposition is a literal part of the moral life of the individual who has undergone the change of heart. But the Son of God symbolizes the disposition to good. That is, we think in relevant respects about our (literally) changed disposition in the same way we would think about such a morally perfect human individual. This makes the Son of God an analogue of our disposition, and as the symbol of our disposition, it provides a way we can represent the object of our Belief – that is, our practical assent to the possibility of our end.

Third, that end expressed in religious terms is becoming *well-pleasing to God*. Our changed disposition and the striving for moral progress it brings about are, once again, a literal part of our moral life. But the end of this striving is represented symbolically in religious terms as God’s acceptance of us. The Son of God presents symbolically what God accepts as well-pleasing to him. In other words, the thoughts we have about the ideal individual enable us to represent in symbolic (religious) terms how the change of heart relates to God’s commands

and the moral individual's hopeful aspiration to be in some way adequate to them, despite having begun from a propensity to evil.

God is thought of as accepting the changed heart or new disposition in the same way God would accept such a morally perfect human individual. This counts as atonement for past evil and makes the human being morally receptive of God's mercy. The Son of God is thought of as a *vicarious substitute*. Treating the ideal individual as a *symbol*, we think in the same way about this individual, and the new disposition that is the outcome of our change of heart, as we would think about someone whose sufferings expiate past deeds of wrongdoing. The Son of God, thought of as a perfect individual, is also thought of symbolically in the same way we would think of an *advocate* who successfully pleads our case for mercy in a just legal proceeding before a holy and righteous but also beneficent and merciful divine judge.

There is, of course, also a possible *fourth* level of symbolism implicit in Kant's account. Those who believe in revealed Christianity might also think of their revealed doctrines belonging to the outer circle of faith as symbols for what is presented in the inner circle of rational religion. For example, if Christians believe that the Son of God was a historical person (Jesus of Nazareth), and that the saving work symbolized by the ideal of humanity was performed by Jesus Christ through his crucifixion and resurrection, then this might be a way in which the doctrines of revealed religion could be seen as "leading back" to those of rational religion, thus offering confirmation of the hypothesis to be tested in Kant's second experiment (R 6:12).

5.5 FAITH IN THE REALITY OF THE SON OF GOD

If through religious symbolism we think of God as a morally perfect being who commands our duties, then in considering ourselves as subjects of these commands, we must measure ourselves and our actions by whether they are in conformity with the divine will or (as Kant puts it in the words of scripture) by whether we can regard ourselves as *well-pleasing to God* (Mark 1:11, Matthew 3:17, Luke 3:22). The task is therefore to ask whether, as flawed moral agents beginning from a radical propensity to evil, we can make ourselves *well-pleasing to God*. The Son of God offers us a way of putting this question because it symbolizes, in the form of an individual, finite will, a moral agent who is well-pleasing to God. Kant emphasizes that we cannot think of ourselves as emulating the ideal because its purity of will is achieved innately and without effort (R 6:64). The Son of God can, however, serve as a symbol for the purity of *disposition* to which a human being aspires and even hopes to attain by undergoing a moral revolution or change of heart.

The symbol of the Son of God thus enables us to pose the question, Might we think of ourselves as capable, through the change of heart, of becoming human beings well-pleasing to God? Since God is represented symbolically as a morally perfect will, this is a symbolic way of posing the question whether we can be morally justifiable in accepting *ourselves* for the sake of the change of heart and the pure

moral disposition. This same question can be posed, again symbolically, by inquiring into the “objective reality of the ideal of humanity” (R 6:62). Or as Kant also poses the question, Is it possible through the change of heart for us to “appropriate” the ideal of humanity and make it comprehensible how this ideal, and the pure disposition it symbolizes, can turn us into human beings well-pleasing to God (R 6:66)?

Someone might think that on Kantian grounds, the question can be answered affirmatively and in literal, moral terms, based simply on the principle “Ought implies Can.” For if we *ought* to make ourselves well-pleasing to God, then it follows that we must *be able* to do this. But that thought misrepresents the situation. “Ought implies Can” entails that we can now do as we ought and set the ends we ought to set. If we ought to undergo the change of heart, and set the end of becoming well-pleasing to God, then we can set this end. But “Ought implies Can” does not entail that we can actually achieve the ends we set when we do as we ought: specifically, the end of being forgiven by God and being well-pleasing to him. The issue raised here is of the same kind as the one involved in the moral argument for God’s existence: we set the end of becoming well-pleasing to God. As with the highest good, “Ought implies Can” does not entail that the end is possible of attainment. These are rather matters for moral Belief, perhaps of a “doubtful faith” (*Zweifelglaube*) or a rational *hope* that sustains commitment and effort in the face of uncertainty.

Someone might suppose that we can be assured of this possibility trivially through the thought that God, represented as beneficent and merciful, would of course find us well-pleasing to him. But this misses an important point. We are not supposed to be thinking of God anthropomorphically, whether as cruel and vengeful or as permissive and lenient. God must be represented symbolically as a *morally perfect will*. The real question is whether a perfect will would find us well-pleasing. The answer to that question is by no means self-evident. A positive answer must remain for us an object of hope or at most Belief, not of confident certainty. Kant poses this question symbolically by asking whether the ideal of humanity or the Son of God has *reality*. The ideal is *real* if the Son of God can make us well-pleasing to God. This is the symbolic way of representing the thought that God in his mercy might find us well-pleasing despite our past misdeeds and continuing propensity to evil, for the sake of the good disposition that is the outcome of the change of heart.

5.6 THE THREE DIFFICULTIES

Kant gives substance to the question of the reality of the Son of God by raising three “difficulties that stand in the way of the reality of this idea” (R 6:66). To Believe or hope that we can be well-pleasing to God is to have faith in the reality of the Son of God.

The Son of God represents holiness of will or the “prototype of the moral disposition in its entire purity” (R 6:61). It therefore symbolizes the good

disposition acquired through the change of heart. We think of this disposition symbolically as the Son of God, a human individual represented as well-pleasing to God. To have faith in the reality of this idea, therefore, is the same as having faith that our own pure disposition, our change of heart, will make us well-pleasing to God. Our striving to become well-pleasing to God therefore rationally requires assent to its possibility, which Kant formulates as “*practical faith in this Son of God*” (R 6:62).

It may help us to understand the issue posed by faith in the reality of the Son of God if we realize that this is the way Kant thinks we can present to the religion of reason the same issue of faith that Christians, for their revealed religion, can pose in symbolic terms by asking about the reality of the person of Jesus Christ and the effectiveness in our lives of his saving work. A Christian believes that Jesus Christ really existed at a past time and place. This belief, as Kant would understand it, symbolizes the rational religious Belief, or at least the hope, that the change of heart has actually occurred in us. The Christian needs also to believe, or at least hope, that through his crucifixion and resurrection, Christ has saved repentant sinners and made them acceptable to God. The change of heart is symbolized as Christ in us acting for good and atoning for our past sins. This corresponds to the rational religious Belief, or at least the hope, that the good disposition achieved through the change of heart will make us well-pleasing to God. The first two of Kant’s three “difficulties” concern the first issue of faith; the third difficulty, which Kant tells us is the greatest, concerns the second.

The *first* difficulty. Kant poses the first difficulty in the following way: “How can this disposition [the good disposition resulting from the change of heart] count for the deed itself, when this deed is every time (not generally, but at each instant) defective?” (R 6:67). We might think that this difficulty is the same as the third – namely, whether, since I began from a propensity to evil, the change of heart all by itself can make me well-pleasing to God. But this would be a misunderstanding.⁷ Kant says explicitly that they are different (R 6:67 n). The issue is not whether the change of heart can make me well-pleasing to God but whether the pure disposition acquired through the hoped-for change of heart is genuine. Kant is distinguishing between saying that our changed disposition is “generally” defective and saying that it is defective “at each instant.”⁸ The first difficulty therefore has to do with the fact that because we

⁷ It is a misunderstanding found in Pasternack (2014), pp. 142–153.

⁸ In the literature on the *Religion*, this is sometimes presented as the “two perspectives” problem. It assumes a supposedly Kantian transcendent metaphysics of timeless noumenal selves and then suggests (to my mind, perversely) that God’s timeless perspective represents to him something that we situated in time know *not* to be true – as if God’s timelessness “solves” the difficulty simply by blinding God to our moral imperfection. That would obviously be nonsense. Or perhaps the basic error here is the same as that of Prichard and Ross, supposing that the duty to act from duty is a perfect duty pertaining to each action in time. For other examples of the error, see Michalson (1990), pp. 103ff; and Pasternack (2014), pp. 150–153. Perhaps the origin of the misreading in the late twentieth century is to be found in Silber (1960), pp. xcvi–ciii. The entire controversy is discussed better by Palmquist (2000); and by Chignell (2016), pp. 110–111.

experience our life within time, our moral striving is always incomplete. At each moment our disposition will be defective, hence necessarily incapable of “counting for the deed itself.”

The change of heart can be manifested solely in a constant moral progress toward holiness of will, which at each instant of our lives remains defective and always less than holy. “The law says: ‘Be ye Holy (in the conduct of your lives) as your Father in Heaven is holy’ [Mt 5:48, cf. Lv 11:44, 1 Pt 1:16] ... The distance between the goodness we ought to effect in ourselves and the evil from which we start is, however, infinite” (R 6:66). Even if our life-conduct *thus far* were to confirm that the change of heart has taken place, this always leaves us at every instant short of moral perfection. We fully satisfy the duty to act from duty in its perfect or qualitative aspect through the change of heart, but we can never confirm that we have done so.

To the “first difficulty,” Kant says,

The solution rests on the following: According to our mode of estimation, who are unavoidably restricted to temporal conditions in our conceptions of the relationship of cause to effect, the deed, as a continuous advance *in infinitum* from a defective good to something better, always remains defective, so that we are bound to consider the good as it appears to us, i.e. according to the *deed*, as *at each instant* inadequate to the holy law. But because of the *disposition* from which it derives and which transcends the senses, we can think of the infinite progression of the good toward conformity to the law as being judged by him who scrutinizes the heart [through pure intellectual intuition] to be a perfected whole even with respect to the deed [the life conduct] (R 6:67).

We should not distract ourselves here with unintelligible metaphysical speculations concerning the supposed timelessness of free choice in a supernatural world. What God’s intellectual intuition knows is our whole life, hence our disposition, *as it really is, in this temporal world*. Hence God can judge the whole of our life, whereas at every instant we see only part of it because the future is still indeterminate and still up to us. What is questionable, therefore, from our limited and temporal human standpoint, is *what* our life-conduct in time presents to God’s infallible knowledge. Because our life-conduct is always imperfect – at best, always progressing and always vulnerable to our self-deception – we are always prey to doubts about this. The issue of faith is whether the disposition has truly undergone a revolution. What we need to hope is that the change of heart has truly occurred, and then also (this will be the third difficulty) we need to ask whether we may hope that is sufficient to make us well-pleasing to God. Overcoming the first difficulty, however, does involve faith in the reality of the Son of God, in the sense that it requires faith, or at least hope, in the face of our always imperfect conduct, that we have truly undergone the moral revolution or change of heart.⁹

⁹ This resolution of the first difficulty is confirmed by a footnote Kant attaches to the passage just quoted: “It must not be overlooked that we do not thereby mean to say that the disposition should serve to compensate for any lack of conformity to duty ... For as regards the compensation for the

It might seem surprising, or even paradoxical, that in response to the first difficulty, my faith is apparently not directed toward God, or the Son of God, but toward myself – and involves God only insofar as his intellectual intuition perceives precisely as they are both the timeless spirit of moral resolve and the progress on which it sets me: have I truly undergone the change of heart? Is my disposition, my fundamental maxim (however imperfectly I follow it in my conduct), truly one of moral purity or holiness? Yet this is still faith in the reality of the Son of God because the Son of God is the ideal that symbolizes the changed disposition on which I must base my hope that I am able to become well-pleasing to God. The issue of faith arises because I can never be certain I've undergone the change of heart. The pure disposition is never my actual possession, and it is therefore appropriate to symbolize it as a being distinct from me – the Son of God, through whom I can hope for salvation.

The *second* difficulty. The second difficulty also concerns the reality of the Son of God as symbolizing the change of heart I hope has taken place in me. Unlike the first difficulty, it does not concern the general distinction between the good disposition and its always imperfect expression in my deeds but rather “the assurance of the reality and constancy of a disposition that always advances in goodness (and never falters from it)” (R 6:67). Even if we can hope that our conduct *thus far* exhibits moral progress, we might well have reason to doubt that we will sustain this progress throughout the rest of our lives. We are constantly vulnerable to self-deception about who we are, which shows itself in the propensity to evil. Good deeds might seem to offer me confirmation of the hope that I have undergone the change of heart, but my occasional moral transgressions might cause me to doubt that I have or perhaps even tempt me to despair. Faith in the reality of the Son of God is what I need in response to this doubt and this temptation. If the first difficulty concerns the doubtfulness of the change of heart from the perspective of its timeless quality, the second difficulty concerns its doubtfulness from the standpoint of its manifestation in time, my always incomplete life as a free being and the lack of determinacy about who I am and what I will make of myself before I die.¹⁰

In the second difficulty, Kant is worried that self-deceiving moral agents will become complacent in the delusion that God has effected a moral conversion in their hearts and cease to worry about the occasional lapses in their conduct (R

transgressions incurred in this progression, we shall consider it in connection with the solution to the *third* difficulty” (R 6:67 n). God’s intellectual intuition of our disposition “makes up only for the deficiency which is in principle inseparable from the existence of a temporal being, never to be able to become quite fully what he has in mind [*im Begriffe*]” (R 6:67 n). This presupposes, however, what may (and even should) always remain doubtful to us – namely, that our disposition has in fact undergone a moral revolution.

¹⁰ Depicting my disposition metaphysically, as an actual thing hidden in a timeless noumenal self, is to reify our freedom, which is not a metaphysical *thing* belonging to another world but a *power* we must presuppose we have in *this* world.

6:68). He thinks that such moral lapses should cause them seriously to fear that they “can reasonably entertain no hope of improving” (R 6:69). Thus Kant is opposed to the practice of those clergy who would offer a dying person the comfort of believing that after death, or even during life, their evil deeds may receive divine pardon based solely on their present *wish* that they should (R 6:69–71, n).

We cannot base [any] confidence upon an immediate consciousness of the immutability of our disposition, . . . but must at best infer it from the consequence it has on our conduct of life. And since our inference is drawn from perceptions that are only appearances of a good or bad disposition, our inference never reveals with any certainty especially the *strength* of the disposition, least of all when, in the face of an impending death, we think we have improved ours. . . . The unavoidable consequence of a rational estimate of our moral state is a feeling of hopelessness (which, however, human nature itself . . . takes care that it does not turn into wild despair) (R 6:71).

Kant’s conclusion, of course, is not that anyone, even a wicked person, should give way to such despair. Still less is it that the moral or religious counselor should cruelly deprive such a person of all hope and consolation. Rather, the doubts and worries that should be present to all of us, at every stage of our lives, ought to “rouse conscience to judgment, to make a break with evil so far as is possible” (R 6:69). Even on my deathbed, what should occupy my thoughts is how to do what good I can still do. The twin enemies of moral striving, as always, are complacency and despair.

The *third* “and apparently the greatest” difficulty. Kant states this difficulty as follows: “Whatever his state in the acquisition of a good disposition, and indeed, however steadfastly he has persevered in such a disposition in a life-conduct conformable to it, he nevertheless *started from evil*, and this is a guilt, blame or indebtedness (*Verschuldung*) which it is impossible for him to wipe out” (R 6:72). The question now is whether, in light of my past sins, even a pure disposition and a change of heart can make me well-pleasing to a morally perfect God. How can I make this possibility comprehensible?

Before offering his own solution to the third difficulty, Kant first takes up and rejects a series of misleading thoughts that might seem to cancel the difficulty or at least lessen it:

First misleading thought: “He cannot regard the fact that after his change of heart he has not incurred new debts (*Schulden*) as equivalent to his having paid off the old ones” (R 6:72). This is all the more true since the change of heart is not literally a datable event in a person’s life, but only a way of representing a resolve. And at any point in time, however sincere the resolve, the person’s past life-conduct will continue to be imperfect and involve transgressions of duty.

Second misleading thought: “Nor can he produce in the future conduct of a good life a surplus over and above what he is under obligation to perform each time; for his duty at each instant is to do all the good in

his power” (R 6:72). This last claim could easily be misunderstood as saying that at each instant we have some narrow or strict duty that we must either fulfill or else incur blame. But as we have seen, that would be fundamentally inconsistent with Kant’s thesis that many of our duties are imperfect and that such a view of our duties could be taken only by a “fantastically virtuous person” who would “turn the government of virtue into a tyranny” (MS 6:409). Kant’s point is better understood if we realize that although some actions are meritorious, and go beyond what we need to do merely to avoid incurring blame, no *new* merit we acquire could serve to earn us *retroactive* permission to have transgressed duty. It cannot remove the guilt for past misdeeds or a past bad heart.

Third misleading thought: The new disposition to good that is the outcome of the change of heart somehow atones for our past guilt when we acted under the influence of the propensity to evil. This thought is closely associated with a common interpretation of the Christian doctrine of redemption through Christ’s crucifixion. The German words *Schuld* and *Verschuldung* can refer either to blame or to guilt, in cases of wrongdoing or else to financial debt or indebtedness. This thought could be interpreted in the latter way, as if our sin were a monetary debt that Christ pays off on our behalf through his passion and death. However, Kant rejects it as an admissible interpretation of the way we might respond to the third difficulty. For, he says, our radical evil is to be understood as our “original guilt (*Schuld*) [that] cannot be erased by anyone else. For it is not a transmissible liability, which can be made over to somebody else, in the manner of a monetary debt (where it is all the same to the creditor whether the debtor pays up or somebody else pays for him)” (R 6:72).

Kant clearly prefers an interpretation of the symbolism – whether that of revealed Christianity or of the religion of reason that we are now exploring – that interprets *Schuld* as guilt and further understands guilt as *liability to punishment*. This point also accounts for Kant’s careful avoidance of the word “forgiveness” in his resolution of the third difficulty because he takes that term to connote the forgiveness of a debt – as by a creditor who might have the authority to collect it or not. After presenting Kant’s solution to the difficulty, I will suggest that whether he uses the word or not, it amounts to an account involving divine *forgiveness* – not forgiveness of a debt but forgiveness in a different (but to us familiar) sense of the term, which we will explore in §§ 7–8 of this chapter and §§ 1–3 of the next.

Finally, before offering his solution, Kant adds to these three misleading thoughts yet a *fourth* mistaken thought that offers yet another false resolution to the problem. The guilt we incur from our maxim of radical evil is *infinite* – not because the being it injures is an infinite being but because it is a guilt found in our disposition (our *bad heart*), not merely in our deeds and omissions. Therefore, the punishment deserved by it is also *infinite* (R 6:72). We cannot suppose that whatever ill may befall us as a consequence of our evil disposition, it could ever

bring us back again to a state where we are beyond reproach – as might happen in the case of a criminal who is released from legal punishment after serving the whole of a just sentence (cf. MS 6:238, 332–333). Our past moral guilt can never be expiated or atoned for in this way, though Kant's solution to the third difficulty will be seen to appeal to a different kind of symbolic atonement.

5.7 GOD'S DECREE OF GRACE

In rational religion we are thinking about our moral duties *religiously* – that is, we are representing them as divine commands. We are representing the radical evil in our nature as disobedience of these commands, therefore, as *sin* (R 6:72) – as a wrong committed specifically against God. Our aim as moral beings, also represented in the same religious terms, is to become well-pleasing to God, despite having wronged him. The third difficulty lies in the fact that there is nothing we can do to erase our past guilt in God's sight or compensate for it. Yet in order to hope that we can become human beings well-pleasing to God, we must be able to hope that things may somehow be made right between ourselves and God. This is a symbolic way of raising the question whether, consistent with rational moral requirements, we can ever make things right between us and *ourselves*. That's a very real question, which we might not take seriously enough if we did not symbolize it in religious terms. Religion enables us to present the question to ourselves symbolically as our second-person relation to a morally perfect divine being, who stands to us in the relation of legislator and also as moral judge. As Adams (2018), p. xvi, emphasizes, our moral striving is not to escape the punishment of sin but to become *well-pleasing to God*, despite our past guilt, which still burdens us. Our task in resolving the third difficulty is to conceive whether and how acceptance by God might nevertheless be possible.

This task can be seen as a way of presenting symbolically the moral constraints on the possibility that God might find us acceptable for the sake of the change of heart. Again, we should review Kant's correct dismissal of some false answers. Meritorious actions in the future can provide us neither with a license to commit new sins nor a way of erasing the guilt for old ones. The liability incurred by sin is not a transmissible one, which might be paid off by someone else, or forgiven as a creditor may forgive a monetary debt. For God to "forgive" our sins in this (monetary) sense would be in effect to suppose that God either forgets we had wronged him or that God now treats our past wrongdoing as morally of no importance – as a rich creditor might choose to ignore a debt to him so small as to be beneath his notice. That would be in effect for God merely to *condone* our wrongdoing. For us even to wish for God's good pleasure on these terms would also be immoral: it would be a further wrong on our part, a further sin. Moreover, we cannot suppose that we might do anything that would morally *require* God to find us well-pleasing. Nothing we could possibly do would impose on God a *duty* either to declare us not guilty or to treat us as not guilty.

The considerations just cited impose on any solution to the third difficulty a set of fairly precise moral constraints. A solution to the difficulty requires that we do, and also undergo, something that makes it morally *acceptable* for us to hope that God might cease blaming us as we have merited by our sin, yet without in any way morally *requiring* of God that he must find us well-pleasing.

Atonement. Kant symbolizes the “satisfaction of divine justice” involved here in terms of the transition from the “old man” to the “new man” represented by the change of heart. The old man (our evil disposition) *must* satisfy divine justice but also *cannot*. The new man (our good disposition), however, is not deserving of punishment. Kant infers that just punishment cannot occur either before the change of heart or after it, but only in the act of conversion itself (R 6:73). “The emergence from the corrupted disposition into the good is in itself already sacrifice (as ‘the death of the old man’) . . . and entrance into a long train of life’s ills which the new man undertakes in the disposition of the Son of God, that is, simply for the sake of the good, yet are still fitting punishment for someone else, namely the old man” (R 6:74). The new man thereby offers himself as a “vicarious substitute” (*Stellvertreter*) for the old man and willingly bears the punishment due to him, at least “for all those who believe (practically) in him.” In this way, the new man is also the redeemer (*Erlöser*) of the converted human being and also the “advocate” (*Sachverwalter*) who pleads the human being’s case before the bar of divine justice (R 6:74).

Here Kant is speaking religiously, hence he is employing symbols – even some symbols obviously drawn from Christianity. But it is all too easy to misinterpret Kant’s symbols. Although Kant speaks here of “rendering satisfaction to divine justice,” there is no possibility of a just or adequate punishment. Nor does Kant think that the ills of life are to be seen as punishments for moral evil; he even goes out of his way to deny this (R 6:74 n). What is most noteworthy, I think, is that the “new man” is described as “undertaking” the “long train of life’s ills” – that is, entering into them *willingly*, merely “for the sake of the good.” Literally speaking, this could not be a case of someone willing their own punishment, since Kant thinks the very concept of doing that is self-contradictory (MS 6:335). We will return in Chapter 6, §§ 2–3 to the question how this “vicarious substitute” is to be understood. For now we see only that it cannot be understood as one person being punished for the deeds of another or one person paying off a debt (conceived as if in monetary terms) that was incurred by another. For now what we will try to do is understand the outcome of this symbolic atonement and advocacy in these terms:

Here, then, is that surplus over the merit from works for which we felt the need earlier, one which is imputed to us by grace (*Gnade*).^{*} For what in our earthly life (and perhaps even in all future times and in all worlds) is always only a mere *becoming* (namely, our being a human being well-pleasing to God) is imputed to us as if we already possessed it here in full. And to this we indeed have no rightful claim (according to the empirical cognition we have of ourselves) . . . It is always therefore only a decree of grace when there is a dismissal of all our responsibility for the sake of this good in which we believe (R 6:75–76).

Kant is steering the careful middle course required by the moral conditions of the problem as we have already identified them. God could not be *required* to issue the decree through which we become well-pleasing to him. But God must also not be thought of as violating any requirement of justice in issuing it. In the footnote (*) attached to the above passage, Kant makes it explicit that this is precisely the meaning of the term “grace” (*Gnade*): “Rather, receptivity is all that we, on our part, can attribute to ourselves, whereas a superior’s decision for which the subordinate has no more than a (moral) receptivity is called *grace*” (R 6:75 n). We have no rightful claim (*Rechtsanspruch*) to God’s well-pleasedness – or in legalistic terms, to his “dismissal” (... *entschlagen werden*) of the charges against us. We do have a (moral) receptivity – that is, God is not required to dismiss the claim against us, but it is no infringement of his justice when he does so.

Kant’s position is therefore misrepresented by Wolterstorff and Pasternack when they say (in Wolterstorff’s words) that on Kant’s account “God is morally required to ... wipe out the guilt of our wrongdoing” (Wolterstorff, 1991, p. 51).¹¹ This claim is incompatible with Kant’s explicit account of the meaning of *grace*: a moral *receptivity* that is *not a requirement*. “Although all human beings are guilty of sin, ... the pronouncement of the judge proceeds [not from justice but] from love” (R 6:146 n). Kant’s account in fact fits precisely Wolterstorff’s own formula (which he mistakenly sees Kant as rejecting): the guilty are still *recognized* as guilty, but they are not *treated* as guilty. Kant deliberately avoids the term “forgiveness” in these contexts because he rejects the parallel with the forgiveness of a monetary debt and also when the same term is used to express the thought that “all accounts must be closed at the end of life” (R 6:70 n). The moral structure of Kant’s account of divine grace can be seen to be exactly the same as that of an important moral phenomenon we easily recognize in human life and that *we* call *forgiveness*.

5.8 WHAT IS FORGIVENESS?

Suppose I have done you some wrong. I am therefore subject to your blame and, if circumstances warrant, perhaps even to some form of punishment. Even if neither you nor anyone else is in a position to punish me, my wrongful conduct is such as to be in itself deserving of punishment. *Blame* is precisely the attitude expressing the judgment that, whether or not anyone may rightfully punish the wrongdoer, wrongful conduct is inherently deserving of punishment; *guilt* is the condition in which someone is to blame. Feelings of guilt are not themselves

¹¹ Pasternack (2014, pp. 156–158) says that although we “may not like what Kant has to say, an answer is given that satisfies Kant” (p. 158). Kant should not be depicted as accepting Pasternack’s misinterpretation of him and then be made to bear the responsibility for it, as if Kant had been aware that his doctrine rightly offends his audience but was being impudently callous about it.

punishments (that is why punishments may actually alleviate them) but rather the feeling that one's conduct is in itself *deserving* of punishment.

Now suppose that in this situation I find my guilt an unsatisfactory condition, and I find it distressing to be the just object of your blame. I want to repair our relationship and make things right between us again, if I possibly can. The way I might try to do this is to do something to *atone* for the wrong I did you. In cases of minor wrongs, an apology may suffice, but if the wrong is more serious, I might have to do something more substantial that involves some sacrifice on my part. I might need to offer you some benefit that compensates you in some measure for the wrong I have done you, even if I can never undo the wrong.

Here it is crucial that I offer my atonement willingly and grounded on my repentance for the wrong. It is then up to you to decide whether to accept what I have done with the aim of reconciling us with each other. If you are conciliatory, you may judge that my act of atonement gives me a moral receptivity regarding your merciful decision to cease blaming me. Unless I do something to atone, any decision on your part not to blame me might itself be immoral – if it would constitute an immoral condoning of my wrongdoing. But if I do atone, it may be morally acceptable for you to cease your blame. By ceasing to blame me, you are showing me mercy, and you do it because, in a spirit of mercy, you accept my act of atonement. The atonement does not undo the wrong, or make my act less blamable or me less guilty, but it does put you in a position where it is morally permissible for you to cease blaming me, to cease treating me as guilty. It may complete the reconciliation between us which I initiated through my act of atonement.

Not just anyone who became aware of my wronging you is in a position to do what you have now done. Someone who was in no way wronged by what I did lacks “standing” (in a sense analogous to the term as used in the law) to decide not to blame me. Further, nothing can create a moral *obligation* for you to cease blaming me. The most we can say, I think, is that if I stubbornly refuse to forgive a penitent and remorseful person who atones for a wrong, then I am behaving with unreasonable harshness and show I lack a certain virtue of love, to which (as was said above) Kant gives the name “reconciliability” (*Versöhnlichkeit, placabilitas*) (MS 6:461).

I believe “forgiveness” is obviously the word in our ordinary moral vocabulary for the relationship just described.¹² It is an essentially second-person relation and also a concept quite distinct from that of forgiving a debt. The phenomenon to which the concept refers is of inestimable importance to our moral and even our political life. Kant's resolution of the third difficulty contains all the elements pertaining to “forgiveness” in this familiar and important sense.

¹² A detailed account of forgiveness that makes the points I have just been trying to make is found in Griswold (2007).

When, by taking up the standpoint of religion (including rational religion), I symbolize my duties by regarding them as divine commands, I thereby also symbolically represent a second-person relationship between myself and God. My transgressions of duty then become wrongs I do against God, putting him in a position to blame and even to punish me. We might say, with justification, that the moral issue that concerns us here is our self-relation: whether we can *find ourselves* acceptable. But a religious person might make the observation that the best way to understand our self-relation here is symbolically. I can be sure of finding myself acceptable to myself only if I think of the issue in second-person terms and ask whether I can think of myself as acceptable to God – whether God forgives me.¹³

As Kant depicts it, the change of heart involves not only entry upon a course of life in which my conduct improves but also an act of atonement for the wrong I did to God as long as I was under the sway of the evil maxim. This act is the willingness to accept the ills of life not as a punishment in the sense of a complete satisfaction but instead as an expression of my repentance for the evil I have done and a resolve to do better in future. God is under no obligation to accept my act of atonement, but this act creates a moral receptivity on my part for his decree of grace. Kant describes the result of his argument at this point as a “deduction of the idea of the *justification* of a human being who is indeed guilty but has passed into a disposition well-pleasing to God” (R 6:76).¹⁴

¹³ Hare (1996), chap. 9, gives an account of forgiveness that is in many ways insightful. At least by implication, however, Hare seems to regard his account as incompatible with Kant’s. He erroneously treats the suffering involved in repentance as a kind of self-inflicted punishment (which Kant would never do) (cf. MS 6:335, 438–439, R 6:24 n; and Hare, 1996, pp. 54–60, 202). No doubt the recognition that we are guilty is painful, but that pain is not punishment. Indeed, to suffer punishment would lessen the pain that comes from the guilty awareness that punishment is still due to you. Forgiveness must find a way of healing that pain, yet in a morally acceptable way. We will see in Chapter 6, §3, how Kant thinks the new man’s vicarious atonement for the old man’s evil heart is supposed to do this.

¹⁴ It is sometimes asked, Can’t the wronged person forgive even an unrepentant wrongdoer – choosing to let go of their own bitterness and resentment? Nelson Mandela is reported to have said, as he was released from decades of unjust imprisonment in South Africa, “As I walked out the door that would lead to my freedom, I knew that if I did not leave my bitterness and hatred behind, I would still be in prison.” The account of forgiveness I am offering here can take account of this important thought in either of two ways: (1) There can still be mercy, and forgiveness in a broader sense, when you let go of bitterness, resentment, and blame because this is a need *you* have. As long as this does not involve your simply condoning the continued infliction of wrong on yourself, it is certainly a justifiable response to your situation, perhaps a morally necessary one. Your mercy helps you, but not the wrongdoer, who still bears the burden of guilt for what has been done to you. (2) There is also what I have heard Miranda Fricker describe as “proleptic” forgiveness: when you hope or even have a faith or Belief that the wrongdoer will repent and you can be prepared to forgive in precisely the sense being presented here, that form of mercy, too, may release you from bitterness. Of course this problem arises only from the standpoint of the person wronged, not from that of the sinner, so it is not directly relevant to Kant’s third difficulty, which is a difficulty for us as sinners, not a problem for God.

The term “justification” is of course a technical one in Christian theology. And a “deduction” in Kant’s technical use of the term is an argument that we are entitled to something of which we already have possession – for example, the categories or the moral law. What we possess here is not actual justification but only the *idea* – the pure rational concept – of justification. We may hope for this justification insofar as we have truly undergone the change of heart. But we can never be certain we have undergone the change of heart; and on God’s forgiveness we can never have a rightful claim but at most only a moral receptivity.

The most we actually possess, however, and what is here given a deduction, is an entitlement to a *hope* on practical grounds for God’s forgiveness. In this connection Kant speaks of *hope* rather than faith.¹⁵ For we can never be so confident that we have undergone the change of heart as to assert that we believe we have. And because it is always up to God to decide whether to forgive us, we should also not presume to declare that we *believe* he does.¹⁶

Kant even warns us that this “deduction” has no *positive* use for the conduct of life because it settles only a “speculative question.” Any “comfort and hope” it might offer is present already in the presence in us of the change of heart “as comfort and hope, not as certainty” (R 6:76). Its chief practical use, he says, is in fact only *negative* because it shows that absolution can depend only on our life-conduct, not on mere feelings, penitential emotions, or ceremonial deeds (R 6:76). Kant’s religious symbolism is never aimed at merely offering solace based on God’s supernatural efforts on our behalf. The religion of reason offers us hope and consolation, but only in the context of our genuine effort toward moral improvement.

¹⁵ As has been rightly emphasized by Chignell (2016). It is noteworthy that when discussing justification through God’s grace (R 6:74–76), Kant speaks three times of hope but nowhere of faith or belief (*Glaube*).

¹⁶ Our faith or Belief is rather directed to the *reality of the Son of God* – who symbolizes the pure disposition we will have acquired *if* we have undergone the change of heart, and also the vicarious atonement and the divine mercy we hope for.

Grace and Salvation

To understand Kant's project in the *Religion*, we need keep in mind, first, how *rational morality* relates to *religion* and, second, how the *religion of pure reason* relates to *revealed religion*. Rational morality is the system of our duties, grounded on the moral law, the autonomy of reason, and the dignity of rational nature as end in itself. As we saw in Chapter 2, Kant argues that there are rational grounds, based on practical rather than theoretical reason, for a morally committed person to have faith (or Belief) in God and therefore to be religious. The ground of morality, however, is not religious; you can be a morally good human being without believing in God or representing your duties as divine commands. But recognizing your duties as divine commands makes you *religious*. For Kant religion *goes beyond morality* in the sense that it represents important parts of our moral life symbolically through our relation to God. As was said in Chapter 1: the right way to describe the relation between morality and religion for Kant is to say that religion adds something to morality that enriches it.

Kant distinguishes pure rational religion from revealed religion, based on alleged empirical revelation from God. Kant's two experiments in the *Religion* consist, first, in seeing whether a religion of pure reason is possible and, second, in seeing whether pure rational religion and revealed religion can be regarded as concentric circles – with rational religion being the inner circle and revealed religion the outer circle. The revealed religion Kant chooses for his second experiment is Christianity. The *Religion* is in effect an attempt to vindicate – at least for certain fragments of Christian revelation – the hypothesis that rational religion and revealed religion are mutually consistent and that revealed religion goes beyond what rational religion can justify. For certain fragments of revealed religion, Kant shows that they can be seen to lead back to the religion of pure reason. The vindication of this hypothesis could also be taken to show that these fragments of revealed religion can serve as a second

layer of symbolism through which the symbols of rational religion can be experienced and lived by a religious person. We will see presently that this is precisely Kant's aim.

Kant's audience in the *Religion* consisted largely of orthodox Lutheran Christians. They believed their salvation depended on the saving work of Jesus Christ. They believed Jesus was a historical person whose coming had been prophesied in the Hebrew Scriptures; that he was begotten by the Holy Spirit of the Virgin Mary; that he led a holy and blameless life; that he was unjustly crucified under the Roman governor Pontius Pilate, offering himself as a vicarious sacrifice for the sins of humanity, and then on the third day that he rose from the dead. They believed that without the death and resurrection of Jesus Christ, we would all be justly damned, but if we have faith in Jesus Christ and his saving work, God's grace makes salvation possible for us. These Christian believers were typically also Augustinians about salvation by grace. That is, they believed we are wholly incapable of saving ourselves and that even our first motion of repentance toward God must be due to God's free gift of grace. Kant's experiments in the *Religion* are aimed at showing these Christian believers that their religion can be reconciled with rational morality and that Christianity will benefit if its symbols are interpreted in that way because this interpretation will help it to fulfill its true religious vocation.

In Chapters 3, 4, and 5 we have explored two of Kant's ways of carrying out his two experiments, specifically in relation to doctrines of sin and of salvation by grace. Chapter 3 treated Kant's way of leading the Christian doctrine of original sin back to a version of it that can be accepted by rational morality and by pure rational religion. This is the thesis of the radical propensity to evil. Chapter 4 dealt with the moral agent's response to the evil in human nature: the change of heart. In Chapter 5 we explored the way rational religion can symbolize the moral realities of radical evil and the change of heart as the striving to become well-pleasing to God. In this chapter we will try to understand the way these matters are further symbolized in revealed (Christian) religion and how Kant understands his project as regards three distinct conceptions of divine grace with which Kant was familiar. Their conventional names are: *prevenient grace*, *justifying grace*, and *sanctifying grace*.¹

6.1 JUSTIFYING GRACE AND GOD'S FORGIVENESS

God's forgiveness or "decree of grace," regarded as part of the religion of pure reason, symbolizes our moral acceptability in the face of the guilt we have

¹ Christian theology offers various accounts of grace and various taxonomies of grace. I choose these three terms because they seem best able to capture Christian conceptions that are recognizably at work in Kant's thinking about grace in both rational and revealed religion. For a different taxonomy of grace in Kant, see Mariña (2017).

incurred through having yielded to our self-inflicted propensity to evil. The “third and greatest difficulty” is posed by our past guilt, which we cannot wipe out. The solution to this difficulty is our faith in God’s decree of grace, which can also be expressed as the hope for God’s merciful forgiveness. “Justifying grace” is the Christian theological name for this solution to “the third and greatest difficulty” posed for our faith in the reality of the Son of God.

As we have seen, Belief or hope for God’s forgiveness depends on faith in the reality of the Son of God, symbolizing the good disposition resulting from the change of heart. As we saw in Chapter 5, Nicholas Wolterstorff argues that morality, or at least Kantian morality, can give no satisfactory account of God’s forgiveness.² He portrays Kantian morality as limited to “the requirements of morality – i.e. of rights and obligations” (Wolterstorff, 1991, p. 46). As Wolterstorff conceives it, Kantian morality apparently requires us to treat people precisely as their conduct deserves (neither better nor worse). This means morality will treat them favorably only if they have a valid moral claim or right to be treated favorably. On Wolterstorff’s account, Kantian morality must say that if the guilty are to be spared blame or punishment, then it must be because they have a moral right or claim on such forbearing treatment. For God to forgive, God must have a strict, narrow, or perfect obligation to treat sinners leniently.

As we saw in the previous chapter, however, forgiveness involves something different from that: forgiveness is a free act of mercy to which the merciful are not obligated and of which, therefore, Kantian morality (as Wolterstorff presents it) can give no account. “Forgiveness is not the declaration that the guilty are no longer guilty but the declaration that the guilty will no longer be treated as guilty. Forgiveness, in that sense, is eminently possible. Often when it occurs, morality is transcended. The forgiven have no moral claim on forgiveness; it comes to them as grace” (Wolterstorff, 1991, pp. 51–52).

In this passage Wolterstorff describes forgiveness correctly, but what he says about its relation to morality misrepresents morality and impoverishes forgiveness. For morality, at least as Kant’s theory presents it, is far more than a matter of rights, claims, and obligations. As we saw in Chapter 4, morality for Kant is more fundamentally a matter of *goodness of heart*: love and respect kept in balance (MS 6:448). Even strict obligations are grounded in both love and respect. But goodness of heart goes beyond them, taking the form of wide duties and also of virtues. Ethical duties in general are in their concept wide, not restricted to what others may claim by right (MS 6:390–391). As we have

² Wolterstorff is by no means alone in thinking that forgiveness “transcends morality.” The idea that religion somehow rises above morality, and especially that divine forgiveness does so, is found in a number of thinkers – see Fichte, *Direction to the Blessed Life* (1806, Lectures 9–10); Kierkegaard, *Fear and Trembling* (2006, pp. 46–59). “Forgiveness” is precisely the point of transition in Hegel’s *Phenomenology of Spirit* from “Morality” to “Religion” (1981, ¶¶ 669–671). For a particularly extreme version of this idea, which claims that forgiveness necessarily violates morality, see Silber (1960), pp. cxxxi–cxxxiv.

already seen, for Kant reconcilability or forgivingness (*Versöhnlichkeit, placabilitas*) is a moral virtue, a wide duty of love. There can be no strict duty to forgive and no strict right to forgiveness; God has no duties whatever (G 4:414, 433), and for human beings it is an imperfect or meritorious duty to be “reconcilable” or “forgiving” (*versöhnlich*) (MS 6:460–461). Duties of love, as wide or imperfect duties, involve treating people in ways to which they have no strict right or claim. This is part of morality itself, by no means its violation or its “transcendence.”

To place forgiveness outside (or “beyond”) morality even impoverishes *forgiveness*. To say *only* that “forgiveness is not the declaration that the guilty are no longer guilty, but the declaration that the guilty will no longer be treated as guilty” makes forgiveness indistinguishable from immoral leniency or merely condoning a wrong. John Hare, too, misrepresents Kant’s conception of God’s mercy when he describes it as the “indulgence” of wrongdoing (Hare, 1996, p. 54). Forgiveness is not *merely* “declaring that the guilty will not be treated as guilty” because that also happens when you simply let another get away with wrongdoing. If you let others wrongly mistreat you, or mistreat someone else for whose welfare you are responsible, and you do nothing about it, then you violate a duty to yourself or to the person mistreated (MS 6:434–437). If that’s what it meant to forgive, then forgiveness would indeed be merely the indulgence of wrong and not a good thing but a bad thing. We saw in Chapter 5, §8, that forgiveness has a determinate *moral* structure: there are moral grounds for its mercy, consisting in the wrongdoer’s repentance and atonement. Those forgiven must have a moral receptivity to mercy but can have no right or moral claim to it. Forgiveness is therefore a freely given mercy that is morally permissible but not morally obligatory.

Perhaps people think forgiveness must be a matter of religion rather than rational morality because, as a matter of history, the religious symbolism involved in human atonement and divine mercy, as presented in the major monotheistic faiths, has been a principal source of our moral thinking about forgiveness. The theme of human penitence and divine mercy – Yom Kippur, *teshuva*, *mechila*, salvation through the sinner’s repentance and the crucifixion of Christ, and *ghufrān*, the limitless mercy of Allah – represents what is most appealing about these monotheistic religions. Forgiveness, however, is also part of rational morality, secular as well as religious. Truth and Reconciliation in South Africa has been a *religious* cause for Bishop Desmond Tutu. But it has also been a *nonreligious* moral concern of Albie Sachs and secular Marxists who struggled and sacrificed to end apartheid but also rejected a legacy of hate and recrimination. If God’s “forgiveness,” too, did not conform to the moral structure of forgiveness – especially if it ignored or violated that structure – then it would be an indulgent leniency unbecoming to God’s nature. It would not be what we rightly cherish under the name of forgiveness.

6.2 VICARIOUS ATONEMENT

From only the standpoint of rational morality, the issues we've been discussing are about my relation to myself: *self*-acceptance or *self*-forgiveness. I can forgive myself if I stand to myself in the relation that a person stands to me when I am morally receptive to that person's mercy but cannot demand it as obligatory. As a moral agent, I have a past, in which I have sometimes committed moral wrongdoing. What is deepest and worst, however, is not merely that I have committed misdeeds. It is that I have had what Kant calls a "bad heart." This bad heart is what Kant means when he thinks of the propensity to evil as the maxim of giving nonmoral incentives priority over the moral incentive. In my heart I have not treated humanity with love and respect. I have cared about the wrong things and placed these things ahead of what I ought to have cared about.

If I have resolved to combat this propensity, then I may hope that I have undergone the change of heart, adopted a pure disposition, and put myself on a gradual path of improved conduct in the future. I now ask whether I can forgive myself for my past misdeeds. This makes sense because all my duties can be regarded as grounded on the constraint of my own reason, hence on a fundamental duty to myself (MS 6:418–419). Before the court of my own conscience, I can be accuser, accused, and also judge (MS 6:437–438). In this self-relation I must judge whether my change of heart (if it is real) constitutes the sort of atonement for my past misdeeds that permits me to show mercy to myself and find myself acceptable by forgiving myself, in the sense of forgiveness we have been discussing. All this could take place entirely within rational morality, without thinking of my duties religiously as divine commands or asking whether I can hope that God will forgive me.

Kant's philosophy, however, offers us some good reasons why I might do better to view these same issues *religiously* – that is, symbolically, in terms of my relation to God. If my duty to pursue the highest good gives me reason to Believe in God, then it also offers me grounds for thinking symbolically of my duties as commanded by God and of God as my judge. Kant treats conscience as a relation of me to myself, but he says it is also appropriate to think of conscience as my accountability to God (MS 6:439–440). The issue of self-forgiveness is never merely a matter of what it would please me to think about myself or what would make me feel self-contented. There is reason to fear that if I view the matter only morally, not religiously, I may never think *honestly* about whether I am *justified* in finding myself acceptable. I would be more likely to settle this issue fairly, being neither too lenient with myself nor too hard on myself, if I present the matter symbolically, thinking of my judge as a supremely perfect will to whom my duties are owed. God is such a being; his judgment reflects perfect holiness of will, perfect justice, but also beneficence. If I ask whether my change of heart and altered disposition make me well-pleasing to God, then this is how to think *religiously* about the issue of self-acceptance.

It can even be argued that there is reason to think about the issue in specifically *Christian* terms. For what I hope gives me moral receptivity for forgiveness is the pure disposition resulting from the change of heart. But whether I have undergone the change of heart is unknown to me; I think of it as known to God because I represent God as seeing and wisely judging the whole course of my life. I can hope I am better than my worst self but only if I honestly face up to my past. I can then represent the pure disposition for which I hope symbolically – as a being separate from me – the Son of God or the “new man.” I hope I may partake of the righteousness of the Son of God if I have truly undergone the change of heart. My hope for forgiveness can be seen to depend, symbolically expressed, on the *reality* of the Son of God or the “new man.”

Kant’s solution to the third difficulty – the greatest difficulty encountered in doing this, was that my past guilt could be thought of as atoned for vicariously by the “new man” – the good disposition, following the change of heart – taking on the ills of life on behalf of the “old man” – the same “physical” person but another “moral person” (R 6:73–75). The acceptance by God of this vicarious sacrifice by this “Son of God” is expressed in God’s decree of grace. The ideal of humanity or the Son of God symbolizes the pure or holy disposition I adopt through the change of heart. It is how I can think of myself as *aton*ing for my past, giving me a *moral receptivity* to (but never a strict moral claim on) God’s grace, which makes me well-pleasing in his sight.

I can distinguish here two conceptions of myself in relation to God’s forgiveness, or two roles I play in seeking God’s mercy. The “old man” is my past, that for which I need to be forgiven. The “new man” is my changed disposition, the outcome of the change of heart, which atones for my past and gives me a moral receptivity to God’s forgiving mercy. My faith, which Kant suggests can never be more than a reasonable hope, is that my changed disposition is real and that God will find me well-pleasing. This is the hope for God’s merciful decree of grace.

We can think of the moral problem we face here as one of releasing ourselves from the burden of a certain kind of *irrational* guilt.³ But when we think of guilt feelings as irrational, we most often think of them as resulting from some mental illness or from a backward moral upbringing that has conditioned us to think of

³ Kant holds, of course, that our will is free, not causally determined, and he does not regard repentance or feelings of guilt and remorse for past misdeeds as itself irrational. But some who deny freedom of the will think that all such attitudes toward our past misdeeds are irrational. This was certainly the position of Spinoza (1982, IVP54) and seems also to have been the position of Sidgwick, 1981, p. 71. But even determinists, who think such feelings are irrational, often hold that they are psychologically inevitable and that we must find a philosophical (or religious) way to deal with them. In his earliest writings, when he was still a Spinozist and denied freedom of the will, J. G. Fichte represented Christian doctrines of vicarious atonement and divine forgiveness as a way people burdened by such irrational feelings might symbolize the self-acceptance that could be thought of with full rationality only in terms of human imperfection and the necessity of the natural order (Fichte, 1970 [1790], 5:5–8).

some things as wrong (e.g., sexual things) that enlightened people can rationally see are not wrong. (Kant's writings show him to have been a victim of some of this pathology, which still afflicts both our cruel moral customs and our deplorable politics.) Overcoming that sort of irrational guilt involves correcting our moral errors through reason and then somehow (perhaps with the help of therapy) bringing our feelings into line with our rational judgments.

The irrational guilt relating to self-forgiveness is, however, of a different kind. It may be even more difficult to handle because it is based not on moral error but on moral truth. For the moral burden of our past is real; it is not merely a morbid illusion. We really are guilty of having had a bad heart. The irrationality consists in the way this burden of guilt threatens to pose an obstacle to the very moral improvement through which we might hope to combat our past wrongdoing. We are tempted to fall into an irrational despair about our moral future. Our past weighs us down, making it harder to hope for a true change of heart. The religion of reason now offers us grounds to Believe, or at least to hope, that the change of heart and our pure disposition have atoned for our past and that God forgives us. The pure disposition is symbolized by the Son of God, for whose sake God's mercy can release us from this burden and enable us to get on with our lives as moral beings. Religious symbolism thus offers us a meaningful and emotionally enriched way of relating to our own past, overcoming moral despair through finding us acceptable to ourselves, despite our bad heart in the past.

Kant's position on this issue has recently been discussed, with her characteristic philosophical insight, moral sensitivity, and good sense, by Barbara Herman. When we wrong someone, we incur guilt that can and ideally should be dealt with through our relation to that person – the relation of forgiveness. But, Herman asks, what if the wrong that troubles us is simply the burden of our past bad heart?

Assume there has been no wrong-doing but only wrong-willing. Suppose that as an adult I realize that misplaced anger at a parent was the driving force of many of my adolescent actions. I would feel bad about this even if there had been no bad actions (perhaps I was an obedient but resentful child), and even if I am not now moved by anger. I feel guilt now at not having been a loving child. There is no repair available. My past interferes with my fully inhabiting the present; it is a consequence of being the kind of agent who has a history. One would want release from such a burden (Herman, 2018, p. 226).

This particular story applies only to some of us, perhaps to only a few. But even fewer, and perhaps no one, could honestly pretend that there is nothing at all in their past lives that burdens them morally, nothing that could "interfere with fully inhabiting the present." Herman's acute abstraction of wrong-willing from wrongdoing serves the philosophical purpose of focusing our attention on the right thing: your bad heart, as distinct from whatever bad deeds you might have done or harm you might have caused to others. Her point is that if Kant is right in employing the thesis of radical evil as a part of moral discipline,

something like it must belong to *any* of us because we all are “the kind of agent who has a history,” in Kant’s view a history burdened by our propensity to evil.

Herman emphasizes that religion need not be the only remedy here (she mentions psychotherapy as another). We have seen that Kant agrees because for Kant there is no moral duty to be religious – to regard our duties as divine commands. In the *Religion*, however, Kant is presenting, and even recommending, a specifically *religious* view of our moral predicament and even a recognizably *Christian* view of it. The response to the burden of a bad heart is faith in the reality of the Son of God and hope that it can render us well-pleasing to God.

Let’s attempt an interpretation of the symbolism here: the changed disposition we hope is present in us is thought of as if it were a morally perfect person or the “Son of God” who takes on the burden of our past, as offering a vicarious atonement for our guilt. Then we think of God as merciful, accepting that atonement through his decree of (justifying) grace, so that through this atonement we become well-pleasing to him. We can find ourselves acceptable to ourselves only because we think of God as finding us well-pleasing to him. The moral problem is that my hope for a better future self cannot be justified until I have come to terms with the self I have been. Rational religion poses my problem by symbolizing what I need as a practical faith (or Belief) in the reality of the Son of God, the vicarious atonement of the “new man” for the guilt of the “old man,” and hope for God’s decree of forgiveness. If I can hope to become well-pleasing to God, then I can be acceptable to myself; I can hope to get on with the task of becoming a better person in the future.⁴

Forgiveness and vicarious atonement. The role in this account of *atonement*, especially *vicarious* atonement, needs further discussion. Kant represents the vicarious atonement made by the Son of God as a punishment that can be thought of as occurring *neither before* the change of heart *nor after* it. *Before* it the old man is simply guilty; he cannot atone. *After* it the new man deserves no punishment and so there is nothing for which he needs to atone. Therefore, Kant argues, it can be thought of as occurring only *during* the change of heart itself (R 6:73–75).

The thinking here, especially in its relation to the change of heart, which is not thought of in relation to time, is obviously *symbolic*. The punishment for past guilt, Kant says, consists in the “entrance into a long train of life’s ills which

⁴ A later, even more explicitly Christian version of the very same issue, is present in Kierkegaard’s *The Sickness unto Death* (Kierkegaard, 1980). The “sickness unto death” is despair: not being able to will to be the self that one is – in other words, not being acceptable to oneself. Kierkegaard (or his pseudonym Anti-Climacus) represents the freedom from this sickness as faith in a God, “the power that posited the self,” and in God’s forgiveness. This is not the place to explain why a Kantian is right to reject Kierkegaard’s (antirationalist) paradoxical solution to the problem of moral despair. But see Wood (2019) for a discussion comparing Kant, Fichte, and Kierkegaard on the concepts of evil and selfhood.

the new human being undertakes in the disposition of the Son of God, yet are a fitting punishment for the old human being” (R 6:74). The atonement is *vicarious* because the pure disposition resulting from the change of heart, the ideal of humanity or the Son of God symbolized by the “new man,” atones for the bad heart of the human being afflicted with the propensity to evil, symbolized as the “old man.” We must keep in mind, however, that this punishment is not a full expiation for past guilt: its result is *not* a moral claim on God. There is no obligation on God’s part to find me well-pleasing or any *obligation* on myself to find myself acceptable. What I seek lies beyond justice; what I need is mercy.

Yet the talk of atonement, and especially of vicarious atonement, may be hard for us to accept. We are bound to think of it as some quid pro quo or even some harsh, legalistic economy of retribution. We may even be tempted to see it as a case of one person paying off the debt of another – despite the fact that Kant even deliberately avoids the very word “forgiveness” precisely because it suggested that to him, and also despite the fact that on Kant’s account the two persons (the “old man” and the “new man”) are physically the same person so that the vicariousness of the atonement belongs to the religious symbolism and not the moral reality symbolized. The thought that after the change of heart, the “new man” no longer deserves life’s sufferings may even seem to us to smack of self-pity.⁵

We may successfully resist these repugnant depictions once we see that the whole point of the religious symbolism is in fact precisely the opposite of what we are thinking when we entertain these thoughts. Kant is attempting, namely, to represent something essential to the moral structure of forgiveness, in fact, the very thing that distinguishes it from the immoral condoning of evil. It is not a matter of seeking exoneration in the context of retributive justice, as a criminal justly punished might rightfully claim to rejoin society with the full dignity of a citizen. That’s the mistake Wolterstorff and Pasternack make when they interpret the charge against us as “being dismissed” (*entschlagen werden*) as a case of God’s *no longer finding us guilty*. For of course God would be *required* to exonerate a person *found not guilty*. (Rather: a good God would never have judged such a person guilty in the first place.) In that case, however, God’s decree could *not* be a decree of *grace* – “a superior’s decision to grant a good for which the subordinate has no more than (moral) receptivity” (R 6:75 n, cf. 6:146 n). So this understanding of God’s grace can’t be right as an interpretation of Kant.

⁵ Self-pity is always to some extent dishonest because it involves casting yourself in the role of mere object, a passive victim of bad treatment, and it therefore means you are dodging responsibility for what you are. When people are accused of “playing the victim,” they are accused of displaying self-pity. But such reproaches themselves are often dishonest evasions of responsibility if issued by those who are complicit in the wrongs. There is no self-pity in a wronged person simply demanding justice. That is precisely the role of an agent with dignity.

We should instead look at the situation this way: when I hope to have undergone the change of heart, I am hoping for a disposition that says, *I will henceforth try to do what is right, strive to be a better human being, to treat others with respect and love – and I won't care about what might be due to me in consequence of this new resolve.* This is a spirit that acts generously come what may, and precisely *does not* worry about any compensation, about any quasi-legal claims or quid pro quo. Lack of concern about what might be due to you is precisely what Kant means when he says that the new man “undertakes an entrance into a long train of life's ills.” It is also as far as possible from adopting an attitude of self-pity, since it the attitude of someone who cares only about what they ought to do and not at all about the fortunate or unfortunate outcome for them of what others might do to them. We can think (symbolically) of the moment of the change of heart as one in which the human being looks forward to life in this spirit of moral generosity and yet at the same time looks back at the life led under the burden of the propensity to evil, which has incurred guilt that can never be wiped away but can at most be forgiven. This attitude is the only way that the aspect of morality that demands justice can play a role in the moral structure of forgiveness, where God's mercy goes beyond what anyone can demand based only on what is due by justice. Herman puts it this way:

The “new man” does not and should not regard the ills and sufferings as something *he* deserves; he can welcome them as a test and exercise for his reformed disposition, of value to him now because of the wrong-willing of the person he once was and as a discipline to stay the course. Following the idea of vicarious substitution, taking the ills and sufferings as a gift makes it true that the new man has left behind the calculus of pleasures and pains that were the marks of his old corrupt principle. *And* by analogy with God's grace, the ills and sufferings can serve for the “old man” as punishment – as paying the debt in a currency the old man understands (Herman, 2018, pp. 226–227).

Herman here makes another good point: when we see vicarious atonement as a fussy kind of retributive justice, we are thinking of the situation precisely as the “old man” thinks about it. That is also why we feel justified in resisting, since of course we want to think of it as the new man does. But if we are being honest with ourselves, our resistance to the old man's point of view will not take the form of mere denial. Even as I adopt the disposition of the new man, I must acknowledge that I am the old man too. The old man's perspective is therefore also essential, if only as what is to be overcome and left behind. That is precisely why the moral structure of forgiveness requires the element of *repentance* and even of *vicarious atonement*. This is how we should look at the difference between forgiving mercy and that immoral indulgence of evil that simply “decides not to treat the guilty as guilty.”

6.3 JESUS CHRIST AS THE SON OF GOD

We have seen already why it makes sense to symbolize the pure disposition resulting from the change of heart as the ideal of humanity, the Son of God, or

the “new man” and to see our becoming well-pleasing to God as faith or hope in the reality of this Son of God. In Chapter 5 we observed that Kant never uses the name “Jesus” or the title “Christ” in speaking of the ideal of humanity, despite the fact that he clearly intends some sort of allusion to the Christ ideal, but cites scriptural passages referring to Christ (R 6:60–62, 65 n, 81–82) and speaks repeatedly of the ideal of humanity as the “Son of God” (R 6:60, 61, 62, 66, 74, 76).⁶ Kant’s presentation of the ideal of humanity as the “Son of God” cautiously holds back from committing itself fully to Christian doctrines regarding the supposed historical person of Jesus Christ – his life, deeds, crucifixion, and resurrection. The ideal human being is portrayed as resisting “the greatest possible temptations,” having “taken upon himself sufferings and a most ignominious death” (R 6:61, cf. R 6: 81, 128) and bearing as “*vicarious* substitute the debt of sin” (R 6:74). “He willingly divested himself [of eternal blessedness] for the sake of plainly unworthy individuals” (R 6:64). At the same time, speaking for pure rational religion Kant resists any commitment to the letter of such Christian doctrines as the virgin birth (R 6:63, 80 n), the bodily resurrection (R 6:128–9, n), and even the Son of God appearing in the world at a given time: “The good principle did not descend among humans from heaven at one particular time but from the very beginning of the human race” (R 6:82).

Here Kant seems to be sending us mixed messages about the Christian way of symbolizing the ideal of humanity or the Son of God. He repeats scriptural language verbatim and endorses its claims. At the same time, in the name of rational religion he denies what Christian scriptures would seem to be saying about Jesus Christ – that he was the good principle descending among us *at a given time* in order to save us from sin. The best way to make sense of these contrasting claims is to distinguish the symbolism that belongs to pure rational religion from the symbolism belonging to revealed (Christian) religion. The pure disposition is a literal reality in the life of someone who has undergone the change of heart. The symbol of the pure disposition, when symbolized in pure rational religion, is the end of becoming well-pleasing to God. The Son of God serves as a symbol when by analogy we think of this ideal in the same way we think of the pure disposition. Kant tells us explicitly that “we cannot think of any significant moral worth in the actions of a person without portraying this person or his expression in human guise” (R 6:64–65 n).

How does Kant think this symbolism, involved in the religion of pure reason, relates to doctrines belonging to revealed (Christian) faith? Clearly Kant does not think that a religion based on reason can regard as necessary the acceptance of a faith in a historical person, his actions or sufferings or the supposed supernatural occurrences attending them.

⁶ Some of Kant’s language even suggests the second person of the Trinity as present in the fourth Gospel: “the Word (the *Fiat!*) through which all other things are” (R 6:60) (cf. Jn 1:1–3).

The living faith in the prototype of a humanity well-pleasing to God (the Son of God) refers, in itself, to a moral idea of reason [which] serves for us not only as a guideline but as an incentive as well . . . to good life-conduct. By contrast, faith in this very same prototype *according to its appearance* (faith in the God-man) is not, as empirical (historical) faith, one and the same as the principle as good life-conduct . . . In the appearance of the God-man, the true object of the saving faith is not what in the God-man falls to the senses or can be cognized through experience, but the prototype lying in our reason which we put in him (R 6:119).

It would be a basic misunderstanding of the *Religion*, however, if we took Kant to be *rejecting* the historical narrative of the revealed Christian faith or *excluding* it from true religion. Kant recognizes that this narrative belongs to the outer concentric circle of revealed (Christian) faith. In line with the hypothesis to be confirmed in the *Religion's* second experiment, it is at least consistent with the inner circle of pure rational religion. The success of the experiment consists in leading fragments of the outer circle back to the religion of pure reason. Kant views faith in Jesus Christ as yet a further level or layer of possible symbolism, one no longer religiously required by pure reason yet an option as part of revealed religion. Historically, Kant says, it has served as a “vivid mode of representing things, apparently the only one at the time suited to the common people” (R 6:83). Even at a later time, it can be a way of representing moral realities that a person might accept by taking enlightened responsibility for the symbolism it involves. We can see Kant doing this in his treatment of such doctrines as the Trinity, the Incarnation, and the bodily resurrection of Christ as they are presented in scripture (SF 7:38–40).

Kant is conspicuously cautious in presenting the doctrines of revealed Christianity. He does not intend to trespass on the domain of the biblical theologian (see R 6: 109–114, 163–168, SF 7:23–24, 36–38). The only claim he is making about the doctrines belonging to the wider circle of revealed religion is that, according to the hypothesis he is trying to confirm in his second experiment, they are *compatible* with rational morality and therefore with the narrower circle, the religion of reason. Kant's evident intention is to offer believers in Christian revelation the option of viewing fragments of their revealed faith as ways of symbolizing those very doctrines that Kant is presenting, also symbolically, as parts of the religion of pure reason. He does not presume to force theologians to accept his interpretation of Christian revelation; he is offering it to them as a possible way to reconcile their revealed faith with their own religion of reason.

As we will see in Chapter 7, this is the way Kant hopes Christians will come to think of their revealed religion. For he hopes revealed religion will reform itself gradually in history through the interpretation of its symbols, thereby more fully actualizing its vocation as true religion. “Historical faith,” he says, “lies at the ground of [religious practices associated with it], until we finally begin to regard them as provisional, and we begin to see in them the symbolic representation and the means of furthering a pure faith of religion” (R 6:176).

In the meantime, Christians may live this symbolism through their revealed faith whenever the historical Jesus in whom they believe provides a prototype for and an incentive to good life-conduct.

To say that the life, death, and resurrection of Jesus Christ symbolizes something about the moral change in ourselves is *not* to say that we *reject* what is presented through the symbol. In fact it is to say precisely the opposite of this. For the symbol is what gives us experiential access to what it symbolizes. Christians for whom it is hard to accept the thought that Christ symbolizes my pure disposition should recall this passage from scripture: “I am crucified with Christ: nevertheless I live; yet not I, but Christ liveth in me: and the life which I now live in the flesh I live by the faith of the Son of God, who loved me, and gave himself for me” (Gal 2:20). As Kant would understand it, Paul’s faith that Christ lives in him symbolizes the way the Son of God lives in the human being who has undergone the change of heart.

6.4 SALVATION BY GRACE ALONE

We may distinguish three elements in Kant’s account that correspond to three kinds of divine grace named earlier. So far we have been dealing with “justifying grace.” But before that there is the act through which the sinner turns away from sin and toward God. Kant describes this act variously as a “change of heart” (R 6:47, 66, 76–77), “conversion” (R 6:69 n, 73), and “revolution” (R 6:47–48). One is what Christians call “prevenient grace.”⁷ Christian doctrine holds that God reaches out to us before we can turn to him. This term also suggests that God’s prior action is involved in every form of grace; also in the third element of the Christian doctrine of grace is its account of the process of moral regeneration following upon the conversion from evil to good. This represents God’s aid given to the sinner, enabling the sinner to attain greater righteousness or moral perfection and to perform good works. It is called *sanctifying grace*. This, too, according to Augustinian doctrine, must be due wholly to God.

Kant is well aware of these doctrines and even this terminology. He denies that we can be saved merely through belief in the letter of these revealed

⁷ Robert Adams argues that Kant rejects prevenient grace (Adams, 2018, pp. xxi–xxii), based on Kant’s claims that if “some supernatural co-operation is also needed for [the human being’s] becoming good, or better, then the human being must nonetheless make himself antecedently worthy of receiving it; and he must accept this help (which is no small matter), i.e., he must incorporate this positive increase into his maxim” (R 6:44). Adams’s conclusion is too hasty. For Kant also says, “A human being’s moral improvement is likewise an affair incumbent upon him, and heavenly influences may indeed always co-operate in this improvement, or be deemed necessary to explain its possibility” (R 6:88). Kant here allows that God’s unilateral action might be necessary even for the change of heart. He regards it as unknowable whether this divine action is needed or actually occurs. But he holds we must strive to make ourselves worthy of it antecedent to depending on it. More will be said about this below in this chapter and the next.

doctrines. Faith in them belongs to the outer circle of revealed religion, which is an option available within true religion but not part of a “saving faith.” The point of saying this is that we may believe in them, but we cannot be religiously required, still less morally obligated, to believe in them. But Kant also holds that these elements of revealed faith might be acceptable ways of symbolizing aspects of the change of heart and reversal of disposition that puts the sinner on the road to moral improvement. At times Kant speaks of the pure disposition as involving a “*practical faith in this son of God*” (R 6:62). Regarding *prevenient grace* Kant repeatedly says that divine aid *may be* necessary to explain the possibility of the conversion from evil to good (R 6:88, cf. R 6:44, 117). He denies only that we can comprehend how it occurs and denies also that reason can either confirm or deny either that it *does* occur or that it *must* occur. Such belief, he says, would be at most useful as a means to effecting our moral conversion, through the thought that God offers us his assistance (R 6:43 n, 52–53, 88, 178, 189, SF 7:43–44).

Regarding *justifying grace* we have already seen Kant speaking explicitly of “justification through a righteousness not our own” and of “vicarious satisfaction.” His solution to the “third difficulty” offers an account of these matters as part of the religion of pure reason. Justifying grace is the only part of Christian grace that Kant directly accepts as belonging to the inner circle of rational faith because he thinks it is important for entering on a good course of life that we Believe, or at least hope, that it is possible for us to become well-pleasing to God. In order to Believe in or hope for that, we must Believe, or at least hope, that the evil principle’s claim on us can be discharged. Rational religion denies that we can do this entirely by ourselves; it requires God’s gracious acceptance of us as well-pleasing to him, despite the evil disposition from which we began. Rational religion, without the need for empirical revelation, can provide this Belief or this hope. This is what Kant calls the “deduction” of the “idea of the justification of the human being” through God’s “decree of grace” (R 6:75–76). That is precisely what we have been discussing in Chapter 5, §§ 7–8, and §§ 1–3 of this chapter.

Kant also explicitly mentions third element: *sanctifying grace*, which is the way God helps us to become better after our conversion (R 6:52, 67, 115–120). He even occasionally calls it by its Christian name: “sanctification” (SF 7:43–44, cf. KpV 5:123 n). Here, too, however, as in the case of the *prevenient grace* that might make possible the change of heart, Kant regards the operation of *prevenient* divine grace as theoretically incomprehensible. It can therefore be neither confirmed nor denied by reason. Acceptance of it (Belief in it, or hope for it) is a religious option but cannot be religiously required.

6.5 MORAL PRESUPPOSITIONS OF GRACE

There are two moral principles that guide Kant in his presentation of the doctrine of salvation belonging to the narrower circle of pure rational faith.

They place constraints also on an acceptable interpretation of revealed Christian doctrines if they are to be consistent with reason and therefore if, as Kant's hypothesis supposes, the wider circle of revealed religion is to remain consistent with the narrower circle of pure rational faith.

First: No loss of the moral predisposition. The first principle on which Kant insists is that although the radical propensity to evil "corrupts the ground of all our maxims," it does not involve the destruction of our predisposition to personality or, what is the same thing, remove the rational incentive we have for following the moral law. We have a self-inflicted propensity to invert the order of these incentives, but that by itself presupposes the continued presence of the moral incentive. The maxim of preferring other incentives to the moral one does not render us *incapable* of giving priority to the moral incentive because it is in that capacity that our freedom consists. Kant thus rejects any doctrine of "total depravity" or "bondage of the will" if they deny that we have lost even the freedom to act righteously. If we lost our moral predisposition, then we would cease to exist as moral beings. That would be inconsistent even with the claim that we continue to sin.

Kant's argument is the following: "The restoration of the original predisposition to good in us is not therefore the acquisition of a lost incentive for the good, since we were never able to lose the incentive that consists in the respect for the moral law, and were we to lose it, we would also never be able to regain it" (R 6:46). Thus Kant insists, in presenting the thesis of radical evil, that "there is one thing in our soul which, if we duly fix our eye on it, we cannot cease viewing with the highest wonder, and for which admiration is legitimate and uplifting as well. And that is the original moral predisposition in us, as such" (R 6:49).

This issue is related to what Kant calls "Spener's problem."⁸ Philip Jacob Spener (1635–1705) was a Lutheran theologian and preacher whose work *Pia desideria* was the chief document of Pietism, the sect in which Kant was brought up. In his maturity, as we have said, Kant's relation to Pietism was always troubled, at best ambivalent. In *Conflict of the Faculties*, Kant addresses a famous religious question raised by "the valiant Spener" – namely, "how religious instruction can make us not merely better human beings but *other* human beings" (SF 7:54). Kant rejects both the Pietist and the Moravian accounts of grace relating to this question because they substitute unknowable mystical claims grounded on subjective feelings for any intelligible answer to the problem (SF 7:54–59). The problem is moral self-knowledge, not metaphysical knowledge; what is needed is not information about supposed supernatural realities and divine acts but the self-honesty necessary to choose good over evil. Thoughts about God might symbolize the needed change in a way that facilitates it but cannot replace what we must do.

⁸ SF 7:58. See Hare (1996), pp. 49–53.

It is significant how Kant understands the distinction between making us “better” human beings and making us “other” human beings. To be a “better” human being, he says, would presuppose “that we were already good but only negligent in our degree of goodness” (SF 7:54). This would be impossible as long as our fundamental maxim is evil. The change of heart, however, would make me into *another* human being and not merely a *better* one. For Kant, to be *another* human being is to have undergone a moral revolution, setting oneself on the path to moral improvement. God’s grace might make us *other* human beings in this sense. It would be nonsense, however, to think of God as utterly annihilating our personalities and then replacing us with *different* persons. That would not be a conception of *our* salvation – the salvation of the *same human being* existing before and after being saved.

Second: The imputability of good and evil. Both the evil we have done, and also anything God could do to save us, must result in acts and moral states that are *imputable to us*. That is what Kant means when he says that “the human being must make or have made himself into whatever he is or should become in a moral sense, either good or evil” (R 6:44). In the literature on Kant, this last claim is sometimes characterized as “the Stoic maxim” (not a term Kant ever uses for it).⁹ As Kant understands it, this “maxim” is equivalent to: “For he ought to become a good human being yet cannot be judged *morally* good except on the basis of what can be imputed to him as done by him” (R 6:51). Kant regards this as simply a condition of moral imputability. If God’s grace is capable of making either better or other human beings of us, then it must do so by producing in us states or acts that are *our* deeds, deeds *imputable to us*. If God’s grace cannot produce such states in us, then it cannot make of us either better or other human beings. (Why would we want to place such limits on God’s power?)

Even in ordinary life we recognize that states morally imputable to a person can be at least partly the work of others than the person whose imputable states they are. If a parent or a teacher educates a child (or an adult) so as to make the pupil a better person, then the moral goodness of the pupil would have to be imputable to the pupil. (And we would presumably say analogous things in a case where one person misleads or corrupts another, making the person evil.) An educator would not have succeeded in making the pupil a better person if the educator had turned the pupil into a mere automaton whose actions are now imputable solely and entirely to the educator and not imputable to the pupil. That the pupil (to parrot Kant’s supposedly “Stoic” phrase) “has made herself into whatever she is in a moral sense,” would be entirely consistent with the pupil saying, “Without that upbringing or education, I would not be the person I am, and my parent or teacher deserves the credit for what I am.” Likewise, if God is to be capable of saving us, then God must be able to produce actions and

⁹ See Wolterstorff, (1991), pp. 48–49; cf. Hare (1996), p. 60; and Chignell (2016), p. 116.

states in reformed human beings that are imputable to *those* human beings. *These very human beings* "must have made themselves into whatever they become in a moral sense," even if God's grace is also to be thanked for it, and even if – as regards God's role in it – God is to be given the *entire* credit for it as *its sole cause*. For if the goodness were not imputable to the human being, then God would not have saved *that* human being.

Imputability of good or evil is not like a pie to be cut up and shared or a monetary prize that, when more than one agent might be involved in achieving it, is divided between them so that each receives a diminished share in proportion to what others have contributed. This picture, in fact, gets things exactly backwards. The successful moral teacher seeks to produce in the good pupil a goodness that is *wholly* imputable to the pupil, however much or little credit (even up to the *entire* credit) may *also* be said to go to the teacher. The more the pupil's goodness is imputable *entirely* to the pupil, the more successful the teacher will have been in improving the pupil. The more that imputation of the pupil's goodness is *not diminished* by the teacher's contribution, the more credit the teacher deserves. Accordingly, if God's grace transforms the sinner, it will have been successful only to the extent that the reformed person's good works are *fully* imputable to the person God has saved.

We must be wary of two misleading thoughts here. *First*, in civil law, burdens are typically shared among those who are jointly liable for harm or damages. But it makes no sense to apply this kind of thinking in cases where it is precisely the purpose of one person to *improve* the actions or character of another. *Second*, a person's imputability for an evil deed may be reduced through diminished capacity, as when mental illness makes a person less responsible. But having been influenced to do something, especially to do *good*, does not automatically make the good deed less imputable to you. It would be absurd to ascribe diminished capacity to those influenced to become good (or better) – *especially* to those wholly transformed from evil to good by God's grace. If a person undergoes the change of heart, and exhibits an improved future course of life, and if we think of God as contributing to these efforts, or even as making them possible at all, then God would have failed in this work of salvation if the contribution of grace to the sinner's reform diminished the imputability of the good deeds the reformed person goes on to do.

6.6 GOD'S CAUSALITY AND THE "INCONSISTENT QUARTET"

It is orthodox Augustinian doctrine that God is *totally* responsible for the goodness of the person who is made good by divine grace. Kant's position is that pure rational religion can neither affirm nor deny that doctrine, simply because we cannot theoretically comprehend how, whether, or to what extent God's supernatural grace is in fact necessary for our becoming good. When it comes to the claim that God's acts are the cause of our change of heart or our

moral improvement, Kant holds that reason must be agnostic: we can know nothing about supernatural causes and effects. If there is a supremely good God, who is also supremely powerful, we can be sure that his intentions are of the best and that he has the power to effect what he intends. But it is enthusiasm or fanaticism (*Schwärmerei*) to pretend we could *sense* such effects. This denies not that they can occur but only that we have the capacity to cognize the supernatural. Therefore, the claim that we need divine assistance to effect the change of heart is not part of the inner circle of rational religion, though it may belong to the outer circle of revealed religion. It cannot be part of rational religion because we cannot claim to know it to be either true or false that our becoming good is entirely due to God's grace. But it is at least religiously permissible to hold, as part of revealed religion, that God's grace is necessary to our becoming good and even that God's grace is wholly responsible for our becoming good.

Kant realizes there is a special problem trying to understand how God's causality acting on our will can be consistent with our freedom. As we saw in Chapter 3, his position is that we cannot comprehend (*begreifen*) – or (what is the same thing) form a determinate concept (*Begriff*) of – how God's causality is compatible with our freedom. For the same reason, we also cannot declare the combination logically impossible or self-contradictory. Kant's position on *divine* causality and human freedom is precisely the same as his position on *natural* causality and human freedom. There, too, Kant thinks we cannot establish the reality (or even the real possibility) of how freedom is compatible with the mechanism of nature. We saw in Chapter 3, §2, that it is a serious distortion of Kant's position, and incompatible with the critical philosophy's strictures concerning human thought and cognition, to ascribe to Kant some positive theory about how freedom and natural causality are jointly possible. Kant's position is best understood in this way: We have two concepts of ourselves as agents: a sensible concept, which we employ in comprehending our actions as subject to natural causality, and an intelligible concept, which relates our actions to practical laws of reason. That we can think of our actions in terms of these two distinct concepts shows that it is logically consistent to hold that we are free and therefore indemonstrable that we are not (since in regard to transcendent metaphysics, this would be the only conceivable demonstration). But we can provide no positive metaphysical account of how freedom and natural causality do fit together or even really can fit together. The most we can show is that our freedom does not logically contradict the causality of nature (KrV A557–558/B585–586). We can form no concept of the combination, however, and can gain no insight into even so much as the possibility of freedom: "Freedom can never be comprehended, nor even can insight into it be gained" (G 4:459).¹⁰

¹⁰ Some writers seem unable to adapt to Kant's total agnosticism about the metaphysical relation of freedom to either natural or divine causality. They cannot resist speculating about what might be

Kant's position on divine causality and human freedom is very close to this. The only differences are these: *First*, we can provide a transcendental proof that our actions belong to the causal mechanism of nature (KrV A189–211/B232–256), but we can have no proof either way whether any of them are due to divine causality. At most it might be an incognizable doctrine of revealed religion that such causality is at work in us. *Second*, we must presuppose that we are free in order to think of ourselves as acting or even as judging, whereas we can never be certain, but can at most hope, that we have undergone the change of heart. This is true whether or not the change of heart might have occurred (in part or wholly) through God's causality.

In his natural theology lectures, Kant discusses the concept of a "moral co-operation" (*concursum moralem*) between God and the free human will. He says, "But as regards a *concursum moralem* or God's free cooperation in the free actions of human beings, such a thing cannot be comprehended in the nature of freedom, but at the same time cannot be regarded as impossible" (VpR 28:110). The *Religion's* position is precisely the same:

The concept of a supernatural intervention into our moral though deficient faculty, and even into our not totally purified or at least weak disposition, to satisfy our duty in full – this is a transcendent concept, merely in the idea of whose reality no experience can assure us. – But even to accept it as an idea with a purely practical intent is very risky and hard to reconcile with reason; for what has to be accredited to us as morally good conduct must take place not through foreign influence but through the use of our own powers. Yet its impossibility (that the two may not occur side by side) cannot be proven either, since freedom itself, though not anything supernatural in its concept, remains just as incomprehensible to us regarding its possibility as the supernatural we might want to assume as surrogate for the independent yet deficient determination of freedom (R 6:191).

In the literature on Kant's *Religion*, there are several different positions taken on the question of Kant's attitude toward the proposition that we require divine grace in order to effect the change of heart and the consequent progress toward being better human beings. The position I have been defending, which seems both to be supported by Kant's texts and to be the only position consistent with his broader philosophical doctrines, is that neither this proposition nor its contradictory are knowable by us and both are even incomprehensible to us. Kant's conclusion is that the religion of reason is not committed to accept this proposition but also committed not to deny it. If the proposition is a doctrine of

lurking in the total darkness to which Kant thinks our cognitive incapacity consigns us. Insole (2016), for example, fancies that Kant is entangled in ultimately unsatisfactory reflections on Christian doctrines of divine grace. He thinks Kant *rejected* Christianity; Kant's merely rational mind could not see its way around in total darkness as well as those who abase their intellects before incomprehensible mysteries and humbly submit in mindless awe to what they proudly proclaim they cannot understand.

revealed (Christian) religion, then the religion of reason should regard it as a religiously permissible option but not required for a saving faith.

Pasternack has defended the position that Kant must reject the need for divine grace based solely on the rational moral principle “Ought implies Can.” He writes: “If the Change of Heart is genuinely morally required of us, then Divine aid is not necessary. It may or may not be available, but if we are to hold consistently to the demand that *ought implies can*, then if this inner transformation is something we ought to accomplish, then it is something we can accomplish . . . Kant opposes the view that we need Divine aid in order to undergo a Change of Heart” (Pasternack, 2014, pp. 145–146). Pasternack’s interpretation would force Kant to assert dogmatically the transcendent metaphysical proposition that divine aid is *not necessary* for the change of heart. But the truth of such an assertion would be unknowable by us; at the same time, it would bring Kant into gratuitous conflict with the Lutheran position he is doing his best to accommodate. That interpretation is therefore both groundless and highly uncharitable.¹¹

It also directly conflicts with what Kant explicitly says: “Everyone must do as much as it is in his powers to do; and only then . . . can he hope that what does not lie in his power will be made good by co-operation from above” (R 6:52). This does not deny that “co-operation from above” might be needed for my conversion or for my further moral improvement. Such a denial would claim more than I could ever know. A morally well-disposed person would surely think that *if* God’s aid is needed, then I should hope it will also be available. But I should not think this divine aid relieves me of the responsibility for doing what I can. On the contrary, it suggests that I ought to do my best to make myself worthy of whatever divine assistance I might need.

The Kantian position should thus be stated as follows: “Ought implies Can. We cannot know whether divine aid is necessary for us to do what we ought. Therefore, if the change of heart is morally required (if it is something we *ought* to accomplish), then *if* divine aid is necessary for us to accomplish it, we *can* justifiably claim that this divine aid is available to us, *if* we do what lies in our power to make ourselves worthy of it.” Unlike Pasternack’s interpretation, this is consistent with Kant’s actual views about “Ought implies Can” and also with what he thinks we cannot know about supernatural divine activity in relation to our freedom.

¹¹ Kant begins one conditional sentence with “Gesetzt, zum Gut- oder Besserwerden, sei noch eine übernatürliche Mitwirkung nöthig . . .” (6:44). Di Giovanni translates *Gesetzt* as “Granted . . .” while Pluhar translates it as “Supposing that . . .” Pasternack repeatedly harps on the difference, apparently thinking that it is significant, and that Pluhar’s translation is superior (pp. 144, 145, 146, 251). But the meaning of what Kant has said is quite clear, however it might be translated. In the German, and in both translations, the import of the antecedent clause is simply conditional: “If . . . supernatural co-operation is necessary . . .” Kant of course wants to leave the truth of this antecedent undecided; we can never know whether it is true.

A position diametrically opposed to Pasternack's is apparently taken by John Hare. He thinks Kant's doctrines commit him to the indispensable need for divine aid in effecting the change of heart (Hare, 1996, pp. 60–68). Because Kant holds that the ground of our maxims is corrupt, Hare thinks Kant is forced into the position "that we may have to grant that some supernatural co-operation is necessary if a person is to become good" (Hare, 1996, p. 61). According to Hare, however, this forces Kant into an "antinomy," because Hare also thinks the so-called Stoic maxim entails that a person must become good solely through his own efforts, thus *excluding* any possibility of divine assistance. "What Kant has to do is show that the revolution is possible, and he does this by pointing to the possibility of supernatural assistance. His failure, however, is to show how he can appeal to such assistance given the rest of his theory, and in particular the Stoic maxim" (Hare, 1996, pp. 61–62).

I hope we have already seen clearly enough where Hare goes wrong here. The propensity to evil disposes me to put other incentives ahead of the moral one but does not deprive me of the moral predisposition, enabling me to give priority to the moral incentive; further, Kant thinks that reason can neither require nor preclude my needing God's assistance in order to do this.

This supposed problem has been formulated with clarity by Chignell, in the form of an "inconsistent quartet": four claims to which Kant is allegedly committed but whose conjunction is inconsistent.

1. S is morally responsible for changing the quality of his will from evil to good.
2. S can make the quality of his will good (from "Ought implies Can").
3. If S is morally responsible for making himself good, and S can make himself good, then S is fully ontologically responsible for making himself good (from the "Stoic maxim").
4. S requires God's assistance in becoming good (Chignell, 2016, p. 112).

Prima facie, (1) and (2) are unproblematic, basic Kantian doctrines. (2) even follows from (1). S ought to make himself good, and he must therefore be able to make himself good. But a contradiction suddenly arises if God is required to do something that only S himself can do. For if S cannot make himself good without divine grace, but the "Stoic maxim" precludes S's becoming good through God's assistance, then it follows that S cannot become good after all. Therefore it must be false either that S ought to become good or that S can do what he ought.

To untangle the knot, let's begin with (4). Is Kant committed to it? *No*. "Ought implies Can" entails that we are capable of the change of heart; it remains wholly unknowable by us whether divine aid is needed for this change. Kant does not dogmatically deny (4). He can without inconsistency entertain it as an optional article of revealed (Christian) faith. But he is not committed to (4) on behalf of rational religion.

What about (3)? Kant never uses any phrase like “fully ontologically responsible.” What could it mean? I fear it may presuppose some of the errors about imputability I have tried to expose in the previous section (nonsensically treating imputability like a pie or a monetary prize). We’ve seen that Kant is troubled when he specifically discusses divine concausation regarding free actions. He finds it “hard to reconcile with reason” that what is credited to us should occur through “foreign influence” (*fremden Einfluß*). “Yet [he says] its impossibility cannot be proven either.” To claim it impossible would be to claim more than we can ever know.

Here’s my guess about what would really threaten a contradiction. If God and I are distinct agents, then it would be self-contradictory to say that some action of which I am the sole author is the very same action as some action of which God is the sole author. But do we know that divine concausation or divine grace must involve that self-contradiction? No. God’s agency and yours might be distinct, each imputable to a separate agent, but also cooperate. We can neither affirm nor deny that this is so, but as with freedom and natural causation, it is incomprehensible but not self-contradictory.

It is not the least conceivable how God might concur in our actions despite our freedom, or how he could concur as co-operating cause of our will; for then *eo ipso* we would not be the author of our own actions, or at least not wholly so. Of course this idea of freedom belongs to the intelligible world, and we are acquainted with nothing of it beyond the fact that it exists, so we also do not know the laws by which it is governed. But even if our reason cannot deny the possibility of this *concursus*, it still sees that such an effect would have to be a miracle of the moral world, just as God’s acts of cooperation with occurrences in the sensible world are miracles in the physical world (VpR 28:1106–1107).

We should not let the term “moral miracle” here spook us. Recall how Kant defines “miracle”: “[Miracles are] events in the world, the causes and *effects* of which are absolutely unknown to us and must so remain” (R 6:86). If God’s grace acts with concausation in producing our free acts, we can never know either that this happens or how it happens. The incomprehensibility of divine concausation can be called a *moral* miracle because imputation is a *moral* concept. If the term “fully ontologically responsible” is meant to imply that acts morally imputable to us *cannot* involve divine concausation, then it claims more than we can ever know. That seems sufficient to show that Kant is not committed to (3), whatever the questionable and mysterious phrase “fully ontologically responsible” might be taken to mean. I think this entire issue illustrates the way we invite confusion and error when Kant’s conception of freedom is thought of in terms of a (necessarily incognizable) “ontology” of transcendent (“noumenal”) metaphysics.

Hare claims that Kant is caught in a contradiction because “the revolution of the will both is possible (because obligatory) and impossible (because “the ground of our maxims is corrupt” (Hare, 1996, p. 61).) Hare is exactly half right: “Ought implies Can” does commit Kant to saying that the change of heart is possible, since it is obligatory. Kant affirms this explicitly: “For if the moral

law commands that we *ought* to be better human beings now, it inescapably follows that we must be *capable* of being better human beings” (R 6:50). But Kant could not be more explicit in denying the other half of Hare’s alleged contradiction. The moral incentive belongs to our moral predisposition. If the moral incentive were not present to us, we would not be moral agents. Radical evil consists in a propensity to invert the rational order of moral and nonmoral incentives, but we cannot invert the order unless both incentives are present to us. Even to be corrupt, the change of heart must be possible for us (perhaps with, perhaps without divine aid – something we cannot know either way).

If Hare thinks that from Kant’s assertions we can infer that the change of heart is impossible, then that, too, is clearly false. “Ought implies Can” directly entails its falsity. If Hare’s claim is that Kant is committed to asserting that the change of heart could be possible *only* through divine assistance, then that, too, is false because Kant holds that this is more than we could ever know. Kant’s position is clear: since we ought to make better human beings of ourselves, we should infer that we can. Does this possibility depend on God’s assistance, perhaps become possible only through it? Reason can never answer that question. Prevenient grace, as a condition of the change of heart, or sanctifying grace, as a condition of our moral improvement, do not belong to the inner circle of rational religion, but they may belong to the outer circle of revealed religion. They are logically consistent with anything we know or can rationally Believe. Rational religion can accept them as permissible but should not require their avowal necessary to a saving faith.

6.7 AUGUSTINIANISM ABOUT GRACE

These considerations should make clear Kant’s position regarding Augustinian orthodoxy and the alleged heresies of Pelagianism and semi-Pelagianism. Pelagianism denies that God’s assistance is necessary for our salvation. Semi-Pelagianism holds that for salvation to occur, we must on our own turn toward God, who will then assist us. Contrary to both, Augustinianism says that both the change of heart and further moral progress are wholly dependent on God’s grace.¹² Kant holds that unaided reason can neither affirm nor deny any of these doctrines. They are all religious options, but none is religiously required. None of them belongs to a saving faith. Any of them might be welcome if it serves the ends of religion by symbolizing parts of our moral and religious life in a way that furthers our moral improvement. Kant rejects Augustinianism only if it claims dogmatically to know things we cannot know (or even comprehend) and then requires us to be intolerant of any equally unknowable and incomprehensible alternative view. (Again, all concepts of “heresy” or “anathema” are alien to Kant’s thought about religion.)

¹² For a reconciliation of Kant with Augustinianism on grace, see Mariña (1997).

The Augustinian doctrine of grace is often connected to the doctrine of election, or divine predestination. Kant considers the doctrine that “a heavenly grace should . . . grant assistance to one human being, yet denies it to another, not according to the merit of works, but through some unconditional decree, and elects one part of our race to salvation, the other to eternal reprobation” (R 6:143). Such a doctrine cannot belong to rational religion, but neither can reason absolutely reject it. Kant classifies it as a “mystery”: it “does not yield the concept of a divine justice but must at best be deferred to a wisdom whose rule is an absolute mystery to us” (R 6:143). However, if we were to suppose that, independently of merit, God were to decree the unconditional reprobation of one part of humanity, such a “doctrine of *predestination* presupposes an *immoral* order of nature” (VpR 28:1116).¹³ Kant also notes that since God is not in time, ascribing to him “prevision” and also “predestination” (*predetermining* in time the eternal fate of human beings) would involve a merely anthropomorphic conception of God (R 6:121 n). We can say that Kant is open to doctrines of election and predestination as long as they are confessions that we cannot comprehend the incomprehensible mystery of God’s will. But such doctrines would be anthropomorphic, superstitious, and contrary to all religion, both rational and revealed, if they impiously represent God’s will as patently unjust.

Toward the end of part one of the *Religion*, Kant asserts two things regarding divine grace that might at first seem incompatible.¹⁴ At the very end of part one, Kant says that “we can admit an effect of grace as something incomprehensible but cannot incorporate it into our maxims for either theoretical or practical use” (R 6:53). Only a few pages earlier, however, he entertained conditionally the thought that prevenient supernatural cooperation might be needed for the human being to become good, or to become better, and then concluded: “The human being must nonetheless make himself antecedently worthy of receiving it; and he must accept this help (which is no small matter), i.e. he must incorporate this positive increase of force into his maxim: in this way alone is it possible for the good to be imputed to him, and that he be acknowledged a good human being” (R 6:44). The first passage seems to say that divine grace *cannot* be incorporated into our maxims; the second seems to assert that it *must* be incorporated into them.

¹³ I don’t pretend to know whether Calvinism or Jansenism really hold the chilling doctrine that God creates some human beings simply in order to reprobate them – that is, to condemn them from all eternity to damnation without ever giving them any option to avoid it. I hope they do not hold this. But if they do, and if (per impossibile) they could somehow compel God to act according to their doctrines, then what would seem to follow is that a good God would reprobate all and only Calvinists and Jansenists because none but they would so blaspheme against God’s goodness as to force a supposedly good God to commit such obvious and terrible injustices.

¹⁴ The possible incompatibility was pointed out to me in discussion by Andrew Chignell. I do not know whether he would accept my proposed resolution of it.

The apparent conflict disappears, however, once we look at the context of the two remarks. The later passage considers an “effect of divine grace” as *merely happening* to a person: “not of our doing, but that of another being – that we, therefore, can *come by it* only by *doing nothing*, and this contradicts itself” (R 6:53). Kant considers it self-evident that *doing nothing* (i.e., *inaction*, simply as such) cannot be incorporated into any of our maxims of *action*.¹⁵ The earlier passage, however, was explicitly about *acting*: “making himself into a good human being.” Kant is granting hypothetically (because reason cannot deny it) that “some supernatural cooperation is also needed to his becoming good, or better.” In that case, the person can and should think that he ought to make himself worthy of this help and also think of himself as accepting the divine assistance as something he must incorporate into his maxim. The two passages are therefore entirely consistent.

In part three of the *Religion*, Kant will argue that revealed religion has grounds for interpreting itself as a vehicle for pure rational religion. This last point has two complementary sides. On the one hand, revealed religion will be seen as the vehicle, the outer shell, of the religion of reason, which is the kernel. On the other hand, revealed religion will be seen as historically indispensable. Revelation cannot be simply rejected, bypassed, or supplanted by rational religion. The relation between the two concentric circles will thus always be one of mutual dependency, never mutual exclusion. Apparent antinomies in relating them are to be resolved by progress in the future history of religion. This is why it immediately makes nonsense of Kant’s entire project in the *Religion* when people ask which (revealed) Christian doctrines Kant accepts and which he rejects. The very question itself makes Kant’s position impossible to understand, much less to take seriously.

¹⁵ One familiar religious thought at this point is that we sinners are actively resisting God’s grace, and if God is to save us, then we must cease this resistance. We must therefore “do nothing” in the sense of *making ourselves passive* to God’s grace. But then our choice to cease our sinful resistance would have to count as an *action*; it would be false to describe it simply as *inaction*. Here as elsewhere, the sycophantic disposition to credit God with everything and ourselves with nothing can lead to patent nonsense if we do not restrain our contemptible impulse toward slavish groveling to the degree needed to preserve bare logical self-consistency.

The Ethical Community and the Church

Part one of *Religion* presents an interpretation from the standpoint of reason of the Christian doctrine of original sin. Part two raises, and attempts to solve, difficulties we may have in conceiving the possibility of wiping out the innate guilt that burdens us through a past burdened by the evil propensity. It expounds, again from the standpoint of reason, Christian doctrines involving conversion, the ideal of humanity (or the Son of God), and divine grace. It defends hope for the change of heart or moral conversion and also for God's decree of grace. But the *Religion* has still not asked how the struggle against evil could be carried on effectively. What can we do now to effect the change of heart in ourselves or in others? The *Religion* has not yet tried to say either what human beings must do or what they have in fact done to make progress in the struggle against evil. Part three addresses this issue.

7.1 THE ETHICAL STATE OF NATURE

Kant holds that each of us individually bears the entire responsibility for the propensity to evil in our nature and must bear the guilt when we yield to it. But the *Religion* also argues that we would not be subject to this propensity except for the social condition – the presence around us of other human beings. Only I, not others, am responsible for my evil propensity, but it is the presence of others that awakens in all of us the propensity to place the incentives of self-love and inclination ahead of the moral incentive. As we saw in Chapter 3, the radical propensity to evil is another name for what Kant, borrowing a phrase from Montaigne, calls the “unsociable sociability” of human nature. This social context and origin of the propensity to evil plays little role in Kant's argument until part three, where it becomes crucial in the struggle against evil:

The human being is nevertheless in this perilous state [the radical propensity to evil] through his own fault; hence he is bound at least to apply as much force as he can muster

in order to extricate himself from it. But how? That is the question. – If he searches for the causes and the circumstances that draw him into this danger and keep him there, he can easily convince himself that they do not come his way from his own raw nature, so far as he exists in isolation, but rather from the human beings to whom he stands in relation or association (R 6:93).

Our propensity to evil is not (or not yet) demonstrable, but Kant thinks it plausible enough for us to accept it and “easily convince ourselves” that it arises from our comparative-competitive relations with others. The answer to Kant’s question “But how?” therefore lies in promoting a different kind of sociability. The prospects for combating evil in ourselves would be bleak if each carried on the moral struggle alone and apart from others. Required is an association in which human beings unite to combat together their propensity to evil:

If no means could be found to establish a union which has for its end the prevention of this evil and the promotion of the good in the human being – an enduring and ever-expanding society, solely designed for the preservation of morality by counteracting evil with united forces – however much the individual human being might do to escape from the dominion of this evil, he would still be held in incessant danger of relapsing into it (R 6:94).

In part three, division one, subsection I, Kant conceptualizes our perilous moral condition using an analogy with political life, calling it an “ethical state of nature,” and treats of the social union, the community or commonwealth (*gemeine Wesen*) that is needed if we are to combat evil in ourselves. By “state of nature” Kant means not an imagined past condition but merely any condition of people living together with no common legal order or authority, whether rightful or ethical. He presents the ethical community as analogous to the civil condition or condition of right, in which people found a political state in order to protect their rights from the violence and injustice of others. But the analogy, as Kant takes pains to emphasize, is an imperfect one. If we are to understand his position, we need to remain constantly attentive to both the similarities and differences.

The *similarities* seem to be these: *First*, in both the juridical and the ethical state of nature, “each individual prescribes the law to himself, and there is no external law to which he, along with the others, acknowledges himself to be subject.” *Second*, “in both, each individual is his own judge, and there is no effective public authority with the power to determine legitimately, according to laws, what is in given cases the duty of each individual, and to bring about the universal execution of these laws” (R 6:95). In both cases, when people join a community, they become subject to a *common legislation* under a *public authority*.

The main *differences* between the juridical and ethical community seem also to be two: *first*, juridical legislation admits of rightful external coercive enforcement; but it would be a violation of right for anyone to coerce the fulfillment of ethical laws. The same is true for entry into an ethical community: “It would be a contradiction [*in adjecto*] for the political community to compel its citizens to enter into an ethical community” (R 6:95). *Second*, the ethical community is in its concept universal, encompassing all humanity, while any juridical

commonwealth is always limited in membership to those residing on a certain portion of the earth's surface.¹ The ethical community, however, is universal: "The concept of an ethical community always refers to the ideal of a totality of human beings. And this distinguishes it from the concept of a political community" (R 6:96).

7.2 THE ETHICAL COMMUNITY

The ethical community not only is subject to public laws but also involves publicly valid *judgments* about their application to particular cases. It also assumes a *power* sufficient for the "universal execution" of its laws. Yet it is not immediately clear how these requirements can apply in a wholly voluntary, noncoercive community, where there can be no judicial pronouncements or their coercive enforcement. It is not difficult, however, to understand what Kant has in mind. Since I regard my ethical duties as divine commands (R 6:153), I treat God's apportionment of happiness to worthiness (in the highest good, KpV 5:110–113, 124–132, R 6:5–6) as analogous to the coercive enforcement of juridical duties by a state. Yet I must not treat God's rewarding or punishing justice as the *incentive* to ethical obedience because that would substitute "compulsory" or "hired service" (*Frondienst*) for the unique ethical incentive (R 6:116, 158, cf. ED 8:338–339). Also, when I join an ethical community, I involve myself in relations with others that help me to determine what my ethical duties require in specific cases, and I freely subject myself to the community's judgment about this, at least insofar as I am able to do so with the approval of my conscience.

In division one, subsection II, Kant argues that just as in the juridical state of nature, everyone is subject to unjust attack by others, so in the ethical state of nature "the human being is incessantly attacked by the evil found in himself and in every other as well" (R 6:97). Ethics lays upon our species the general duty of struggling against this evil. "Now, here we have a duty *sui generis*, not of human beings toward human beings but of the human race toward itself" (R 6:97). This is a duty to belong to a community based on a shared end – "the highest good as a good common to all" (R 6:97).

In other words, Kantian ethics is not individualistic but fundamentally communal. This is contained already in the most final and definitive formula of the moral law, the Formula of the Realm of Ends, since this realm is an ideal community of all rational beings united by a system of shared ends (G 4:433).²

¹ Kant holds that a universal world-state would necessarily be "a soulless despotism" and "the graveyard of freedom" (EF 8:367).

² Kantian ethics is justifiably called "individualistic" if what is meant is any of the following: (1) Kant's theory of right is grounded on the individual human being's innate right to freedom (MS 6:239–238); (2) Kant's theory of imputability recognizes no collective guilt; it makes individuals responsible only for their own deeds and volitions (MS 6:227–228); and (3) the fundamental value in Kantian ethics is the dignity of humanity in individual persons as ends in

In the ethical principle, however, this community is merely an ideal, used to identify those laws to which the individuals are subject (they are the laws that, if universally followed, would bring about a combination of ends and the ideal community of rational beings). The ethical community proposed in the *Religion*, however, is supposed to be not a mere ideal but an actual community on earth. Yet it, too, is an *ideal* in the sense that no existing earthly community ever fully lives up to it; it is an idea (pure normative concept) toward which we must orient actual human communities that have been established to struggle against evil.

It is essential to the purpose of the ethical community that it must be grounded on a consciously shared common end (the highest good). For the source of evil is our unsociable sociability or competitiveness, which sets human ends in conflict: in the ethical state of nature, human beings “deviate through their dissensions (*Mißhelligkeiten*) from the common goal of goodness” (R 6:97). The struggle against evil therefore requires a different, contrary kind of sociability, which involves not dissension but unanimity (*Einbelligkeit*), hence the consciously collective action of human beings and the conscious unification of human ends into a universally shared end – the highest good.

In division one, subsections III–IV, Kant argues that a community (*gemeines Wesen*) consists in a group of people united by (1) common or shared ends and (2) a system of common legislation. To constitute the legislation for a community, however, this legislation must be public, seen by members of the community as “the commands of a common legislator” (R 6:98). Kant argues that if the legislation is to be ethical rather than juridical, the people cannot be seen as this legislator because ethical legislation is addressed to the inner disposition rather than to external actions (it commands morality, not mere legality).

This argument would not be cogent if the only problem were *what* the laws command because there is no reason in principle why those who command would, in that capacity, have to be able to know whether the commands are obeyed. But if the commander is also a *legislator* – someone who lays down sanctions (rewards or punishments) for obedience or disobedience (see MS 6:227) – then these laws cannot be justly administered unless the legislator does know the merits of those subject to the legislation. Because the ethical community is united by a shared end (the highest good) and because one component of the highest good is the apportionment of happiness to inner worthiness (which is thought of as a sanction imposed by the legislator), the legislator for the ethical community must be thought of as “one who knows the heart” so that he can “give to each according to the worth of his actions” (R 6:99). The people (*Volk*) or human members of an ethical community, if they were regarded collectively as its legislator, could sanction actions not

themselves (G 4:428–433). But like later great representatives of the German idealist tradition (Fichte, Hegel, Marx), Kant saw clearly that if we treat human community and individual dignity as opposed or competing ends, we misunderstand both ends and we will never achieve either.

according to their morality but at most according to their legality. Therefore, only God can be the legislator for the ethical community, which accordingly is thought of as a “people of God, and indeed in accordance with laws of virtue” (R 6:99).

7.3 ECCLESIASTICAL FAITH

The actual ethical community is not the realm of ends, regarded as a mere ideal (G 4:433–434). In order effectively to combat evil, the ethical community must exist on earth, as a human institution. This means it must come to terms with human weakness, fallibility, and corruption. “The sublime, never fully attainable idea of an ethical community,” however, “is greatly scaled down under human hands” and displays itself as what is called a “church” (*Kirche*). This human institution displays certain restrictions “under the conditions of sensuous human nature. But how could one expect to construct something completely straight from such crooked wood” (R 6:100, cf. I 8:22–23)?

Adopting traditional Christian terminology, Kant distinguishes the “church invisible” – “the mere idea of a union of all upright human beings under direct yet moral divine world-governance” – from the visible church – “the actual union of human beings into a whole that accords with this ideal” (R 6:101). This accordance, however, will always be imperfect and even encumbered by certain essential limitations, grounded in the imperfections of human nature. This, as Kant explains in subsection V, takes the form of a historical faith, which Kant calls “ecclesiastical faith” (*Kirchenglaube*), involving revealed statutory laws of divine service. Subsection IV formulates the idea of a true ethical community, according to the four headings of the table of categories (R 6:101–102):

1. *Quantity*. The ideal ethical community would be *numerically one*, and *universal*, encompassing all human beings. There are a plurality of ecclesiastical faiths, but we should see them all as expressions of a single religion – that is, a single disposition to regard all duties as divine commands.
2. *Quality*. It would be *pure*, measuring human actions solely by their conformity to moral incentives, “cleansed of the nonsense of superstition and the madness of enthusiasm” (R 6:101). As we have seen, *superstition* involves not merely theoretical error but violation of the laws prescribed by our rational understanding (KU 5:294). This means that superstition is not merely a theoretical defect but a moral flaw. *Enthusiasm* (*Schwärmerei*) consists in being persuaded that we can cognize the supersensible, especially that we can empirically cognize the effects of grace or other supernatural operations (R 6:53, 83, 114, 174, KpV 5:84, 123, SF 7:39 n, VpR 28:1109). It, too, is cognitive error grounded on a moral defect.
3. *Relation*. The ethical community would be *free*, both internally, in the relation of the members among themselves, and externally, in relation to the political state. Since ethical laws are essentially noncoercive, an ideal

ethical community would have to be external to the state and it would not itself take the form of a coercive government (or “theocracy” under the aristocratic administration of priests) (R 6:102, cf. 6:100).

4. *Modality*. The constitution of the ethical community would be *unchangeable*. This refers, however, only to the ideal constitution of the community. Kant emphasizes that accidental regulations concerning the administration of the church must change with circumstances, and there can be no permanent set of written laws or creed. Both would be devoid of ethical authority and even stand in the way of enlightened progress in matters of religion (R 6:102, cf. WA 8:38–39). Kant concludes that the church may have

nothing in its principles that resembles a political constitution . . . It could best of all be likened to the constitution of a household (a family), under a common though invisible moral father, whose holy son who knows the father’s will and yet stands in blood relation to all the members of the family, takes its father’s place . . . [the members of the church] therefore honor the father in him and thus enter into a free, universal and enduring union of hearts (R 6:102).

The ethical community in its purity would be grounded solely on ethical laws of pure reason. But as a human institution, subject to human imperfections and limitations, a church displays a certain characteristic weakness of human nature that Kant thinks cannot be directly avoided but must be accepted as part of any earthly institution through which a pure ethical community is to be established.

Human beings cannot conceive their obligation [to God] except as directed to some *service* or other which they must perform for God – wherein what matters is not the intrinsic worth of their actions as much as, rather, that they are performed for God to please him through passive obedience, however morally indifferent the actions might be in themselves (R 6:103).

In other words, human beings have not been able to help thinking of God by analogy with a human potentate or despot, “a great lord of the world [who] has a special need of being *honored* by his subjects and of being *praised* through signs of submissiveness” (R 6:103). An ethical community can come into existence historically, therefore, only in the form of a social body having such statutory laws commanding divine service, which must therefore be given not by pure reason but only empirically, through some alleged revelation of God’s will, as to the manner in which he is to be served, honored, and praised (R 6:103). Mere tradition (the handing down of these statutes orally from each generation to the next), Kant argues, will prove inadequate to preserve them in a stable and unchanged way. Therefore, an ecclesiastical faith will base itself on an alleged revelation written down in a scripture or holy book (R 6:106–107).

For Kant there is only one true religion. But this religion manifests itself – and historically it *must* manifest itself – only through many ecclesiastical faiths. What are called different “religions” and “religious” struggles, “religious” persecutions, and the insistence on “religious” orthodoxy – all these, Kant says, concern only ecclesiastical faiths, not religion itself. It is the collective duty of humanity, part of

their duty to found a true ethical community, to overcome these human deficiencies that lie at the foundation of ecclesiastical faiths and produce the conflicts and historical divergences in merely statutory observances and the superstitions with which revealed faiths are originally encumbered (R 6:107–109). Kant denies that we can ever hope to grasp the pure rational religion directly, either founding a new church based on it alone or arriving at it, so to speak, behind the back of the ecclesiastical faiths. For it is through these faiths alone, Kant thinks, that religious community is accessible to us at all.

We can now understand better why it is important to Kant to conduct his two experiments in the *Religion*, involving two concentric circles: the inner circle of pure rational religion and the wider circle of revealed religion. It is also why Kant hopes to convince his readers that the two circles must stand in the relation he has hypothesized they will stand. They must be at least consistent: there must be no “anathema” by which rational faith condemns any part of revealed faith or revealed faith condemns the religion of pure reason. The wider circle of revealed religion must also lead back, at least in part, by way of an interpretation of its symbolism, to the religion of pure reason. This also shows why Kant cannot be understood as *rejecting* revealed (Christian) religion. He must *accept* it, as the outer circle of religion.

Kant’s *Religion* is addressed simultaneously to its intended audience in two roles: as rational moral agents and also as believers in revealed Christian faith. We must interest ourselves not only in what God demands of us as good life-conduct through the moral law of reason but also in what God can be thought of as demanding of us as members of a real earthly community established to honor and serve him. “The question *How does God will to be honored in a church (as a congregation of God)?* appears unanswerable by mere reason, but to be in need of a statutory legislation proclaimed only through revelation, hence of a historical faith we can call ‘ecclesiastical’ in contradistinction to pure religious faith” (R 6:105).

7.4 THE NEED FOR REVEALED RELIGION

There are two contrasting thoughts here, which stand in some tension, but Kant thinks they can – and even *must* – be reconciled. The *first* thought is that true divine service, that through which God (as moral reason presents him to us) would want to be honored, consists solely in the performance of our ethical duties, symbolically presented in religious terms, but does not consist in obedience to any words of praise, flattery, or obedience to statutory laws proving our cringing submissiveness to God’s arbitrary will (R 6:104–105). The *second* thought is that the rational idea of an ethical community can be realized by human beings on earth and in human history only through institutions involving divine statutory laws and the authority of a written scripture.

Thus, if human beings are to struggle against their radical propensity to evil, they must accept the existence of an essentially flawed institution as the

indispensable vehicle of the struggle. Religion must be subject to a *historical* dynamic, through which the two contrasting thoughts can be brought together. We must accept the existence of churches with statutory laws allegedly revealed empirically through scripture, but then we must struggle to reform these churches so as to bring them gradually into agreement with the *idea* of the ethical community.

This means that in order to fulfill its own proper vocation as religion, ecclesiastical faith and practice cannot remain static. Churches and ecclesiastical faiths must not be a dead weight on humanity, holding it down or holding it back from fulfilling its vocation in history. Part three, division one, subsections VI and VII of the *Religion* explains this historical dynamic. Division two looks at it as it has displayed itself in empirical religious history (especially in the history of Christianity). The historical origin of religion, as Kant understands it, was an attitude of slavish submission to supernatural powers. Its teachings were thought to have been revealed empirically in scriptures handed down by tradition and interpreted by an aristocracy of priests whose chief ambition lay in enslaving the souls of those the society entrusts to their care. This is about as far from a genuine ethical community as can be imagined.

It will help us here to appreciate the way in which this tension in Kant's thinking follows the same pattern we see in other aspects of his philosophy of human history. The starting point for understanding this pattern is the third proposition laid down by Kant in his essay *Idea for a Universal History with a Cosmopolitan Aim*: “Nature has willed that the human being should produce everything that goes beyond the mechanical arrangement of his animal existence entirely out of himself, and participate in no other happiness or perfection than that which he has procured for himself free from instinct through his own reason” (I 8:19). Nietzsche, who loved paradoxes above all else, and especially loved to sharpen them to the uttermost extreme, put the basic thesis of Kant's philosophy of history this way: “All good things were formerly bad things” (Nietzsche, 1967, Third Essay, §9).

What this means is that for Kant human institutions always begin far removed from, even in certain ways directly opposed to, the rational aims that human beings must realize through them. Kant sees this as an endless struggle to free these institutions from their original form and turn them into the very opposite of what they were at the beginning. We may see this pattern in Kant's treatment of political history and even in the history of morality itself.

The political state, in Kant's view, arose in consequence of an economic revolution – the appearance of agriculture, over against the more “primitive” or “uncivilized” hunter-gatherer and pastoral economies. The economic surplus generated by agriculture permitted the growth of towns, and the accumulation of wealth needed to support a military force, that was also needed to protect from pastoral nomads rights of property in land (which were required for agriculture) (MA 8:118–120, cf. KrV A ix). The earliest states were military despotisms. Human political history is the story of the struggle of human beings

to transform a naturally evolved, inherently unjust despotic social form into “a civil society universally administering right” (I 8: 22). The human species is to perfect such a condition of right. Kant regards this as an endless process.

Something analogous holds even regarding morality itself. According to Kant the psychological and social origins of morality lie in imaginatively transformed sexual desire and sexual refusal, which underlie the human aspiration to be valued and at the same time respected by others. This aspiration takes the social form of “decency” or “decorum” (*Anständigkeit*) or “propriety” (*Sittsamkeit*) – the desire to acquire worth in the opinion of others (or at least to avoid their contempt), hence also to confirm our worth in our own eyes, by conforming our behavior to accepted social mores or customs (*Sitten*) (MA 8:112–113, cf. VE Collins 27:300). This is about as far removed as we can imagine from a law of autonomy that we regard as proceeding from our own rational will. But the history of morals is to be a process through which human beings have transformed (and are still transforming) the principles of conduct.

There certainly is something paradoxical in Kant’s view of history. That’s what he means when he says that almost all things human are paradoxical (WA 8:41). But there is nothing anomalous in his idea that in its historical origin, the true ethical community was grounded on the superstitious belief that God is to be pleased not by good life-conduct but by obsequious praise and statutory observances. Kant’s thesis of “historical irony,” as we may call it, is his proposition that an ethical community is possible only through the historical progress of existing churches. It arises from what Kant does not hesitate to call a “peculiar weakness of human nature” (R 6:103). There is, however, an aspect of religion arising from certain parts of Kant’s aesthetics, especially his conceptions of the sublime and its relation to religious symbolism.

7.5 SUBLIMITY AND REVEALED RELIGION

Kant is sometimes reproached with separating morality from everything sensuous, turning it into a lifeless formalism that could never motivate flesh and blood human beings.³ But this is an error, and a caricature. The very purity of the moral law, in his view, gives it emotional power. We should attend to what Kant himself had to say on the topic:

It is utterly mistaken to worry that if [the representation of the moral law] were deprived of everything that the senses can recommend it would then bring with it nothing but cold, lifeless approval and no moving force or emotion. It is exactly the reverse: for where the senses no longer see anything before them, yet the unmistakable and inextinguishable idea of morality remains, . . . it would be . . . necessary to moderate the momentum of an unbounded imagination so as not to let it reach the point of enthusiasm (KU 5:274).

³ For an especially vehement example of this reproach from a feminist perspective, see Schott (1988). But cf. Wood (2008), chap. 13; Herman (2002); and Longuenesse (2017), chap. 8.

The critics, therefore, are entirely mistaken in thinking that a morality that disdains empirical motivation must lack motivating power. Kant is speaking in this passage about the pure moral incentive, which we discussed in Chapter 4, §§ 1–2, as it is expressed in the feeling of the sublime and, in particular, about the role of this powerful feeling in Jewish and Islamic monotheism, such as the Torah’s prohibition on representing the divine in sensuous form and Islam’s prohibition on the representation of humans and animals. Kant’s critics make the same mistake made by ancient idolatrous religions (what Islam calls the *Jahiliyya*) when they underestimated the spiritual power of Hebrew monotheism. Kant’s point is that it is precisely by withdrawing the moral incentive from the limitations of the sensuous that we make possible the most powerful emotional motivation on its behalf. That emotion is found in the sublime, in which we feel something of which we can never have empirical cognition – namely, the moral law’s transcendence of everything merely sensible (KU §§ 27–29).

As we have already seen, Kant thinks religions must make use of this experience in their symbolism of what belongs to the moral life: Thus the Christian conception of Heaven and Hell

is indeed a figurative [*bildlich*] representation, and as such, a stirring [*empörend*] one, yet not any the less philosophically correct in meaning. For it serves to prevent us from thinking of good and evil, the realm of light and the realm of darkness, as bordering on each other and losing themselves in one another by gradual steps . . . but rather [represents] them as separated by an immeasurable gap . . . The total dissimilarity of the basic principles . . . justify this form of representation which, though containing an element of horror [*Schauerhaften*], is nonetheless sublime (R 6:60).

Here we must also recall that for Kant there can be no empirical cognition of the divine (since God is an idea of reason, to which no empirical representation could ever be adequate). Our metaphysical idea of God must therefore be presented symbolically (KU 5:351–354). In the *Religion* Kant repeatedly refers to religious texts purporting to be divine revelation as presenting moral concepts “symbolically” or “schematically” in ways that “further a pure faith of religion” (R 6:176, cf. R 6:111, 136, 171). Kant therefore not only allows it to be permissible to acknowledge a kind of empirical divine revelation but even sees this aspect of ecclesiastical faith as capable of serving the ends of pure moral religion.⁴

What matters to Kant is the correct ordering, and also the correct interpretation, of our thoughts about God. The greatest danger – represented by all ecclesiastical faiths of which he thinks we have any record – is “anthropomorphism.” This

⁴ We have seen that Kant regards religion as essentially symbolic, so that without symbols with emotional power, there could be no religion. Resistance to idolatry therefore misunderstands the problem when it takes the form of *iconoclasm*, or the prohibition on sensible symbols. The only way to avoid idolatry is to accept symbols but avoid literalism – e.g., treating symbols as means of grace (R 6:199) – accepting the responsibility to interpret religious symbols.

religious aberration represents God, following the defects of empirical human nature, as if he were a capricious cosmic tyrant who must be placated and flattered by slavish devotion. Anthropomorphic origins cling to all claims made by such faiths to empirical divine revelation. In order to appropriate revelation for true religion, the proper order of our thoughts should be first to approach such claims from the standpoint of a rational, moral faith and then interpret any proposed empirical revelations as symbolic expressions that can animate and enrich true religious faith. That manner of appropriation should serve as the criterion for accepting alleged empirical revelation in scriptures and also for the hermeneutics applied to them. The same principle, in Kant's view, ought to govern the progressive reform of all ecclesiastical faiths, as they interpret the symbols of their revealed religion in a way that brings it into increasing harmony with pure rational religious faith.

Thus no one can *first* be convinced of the existence of a highest being through any intuition; rational faith must come first, *and then certain appearances or disclosures could at most provide the occasion for investigating whether we are warranted in taking what speaks or presents itself to us to be a Deity*, and thus serve to confirm that faith according to these findings (O 8:143, italics Kant's).

These "appearances or disclosures" always presuppose our rational concept of God. They must always confirm – never supplant or contradict – a moral faith based on reason, which Kant also describes as "inner revelation" (VpR 28:1117).

Nevertheless, Kant does allow the possibility of experiences of revelation and even explicitly identifies them with the purported scriptural revelations of ecclesiastical faiths. He does not *reject* revelation. But revelations are not *cognitions* of God; they are experiences presenting symbolically the contents of practical faith or Belief. The historical necessity of ecclesiastical faiths based on revelation is therefore not due *only* to a weakness of human nature. It is also the only way that beings like ourselves, who are both rational and sensuous, can experience morality's transcendence of the senses and thereby appropriate the experiential and emotional side of moral motivation itself.

7.6 THE INTERPRETATION OF SCRIPTURE

For Kant the very existence of reason depends on something social: free communication (KrV A738/B766). In just the same way, the very existence of religion depends on the existence of churches and ecclesiastical faiths. But in order to be examples of true religion, there must be a process of gradual transformation of ecclesiastical faith according to the idea of a true ethical community. It is essential to the existence of true religion, therefore, that the scriptures of ecclesiastical faiths be subject to continual reinterpretation in a way that promotes the proper end of all religion: "to make better human beings" (R 6:111). For this "we require an interpretation of the revelation we

happen to have, i.e. a thoroughgoing understanding of it in a sense that harmonizes with the universal practical rules of a pure religion of reason" (R 6:110). The pure faith of a rational religion is therefore the "supreme interpreter" of every ecclesiastical faith and the scriptures on which such faiths are based.

The same conclusion follows, as we have seen, from the way in which any scripture could authenticate itself as a possible divine revelation. The concept of God is an idea of reason, to which no experience could ever be adequate. No experience, therefore, could ever provide us with a cognition of God or a cognition that some communication was from God. No writings handed down from tradition could be cognized as an actual divine revelation either from external marks (historical tradition or authentication by priestly authorities) or internal marks (such as their narration of miraculous events). The closest we can ever come to an *authentication* of a purported divine revelation is to decide on rational grounds that its teachings, as far as we are able to determine, harmonize with pure practical reason – with what God *would* teach us if he chose to reveal himself (cf. O 8:143). "For the kind of characteristics that experience can provide can never show us that a revelation is divine; the mark of its divinity (at least the *conditio sine qua non*) is its harmony with what reason pronounces worthy of God" (SF 7:46).

It follows that a necessary condition for regarding a scripture as divine revelation is that it admit of some interpretation according to which it harmonizes with the pure religion of reason. "The only faith that can found a universal church is pure religious faith, for it is a plain rational faith which can be convincingly communicated to everyone, whereas a historical faith, merely based on facts, can extend its influence no farther than the tidings relevant to a judgment on its credibility can reach" (R 6:103, cf. 6:108–109).

There are two ways that this argument might be taken. The *first* way contrasts the merely limited extent to which a historically grounded religious message (e.g., that of the Torah, the Gospels, or the Koran) has at any point been disseminated with the universal extent of purely rational knowledge (such as that of mathematics or the principles of morality). This might invite the rejoinder that of course even such rational knowledge also requires empirical dissemination by a scientific or philosophical tradition. Obviously not all people have read Euclid's *Elements*, Newton's *Principia*, or Kant's *Metaphysics of Morals*. In fact, it is likely that the Holy Bible and the Koran have had a far wider dissemination than any of those sources.

Kant's argument, however, should instead be taken in a *second*, quite different way. Kant can point out that historical claims are grounded on testimony, hence dependent on the extent to which this testimony has been disseminated, while mathematical or moral claims are grounded on reasons that are in principle accessible to all, whatever the historical process through which these reasons might first be made known to people. In that sense, the faith of pure rational religion, even if less widespread than some historical, ecclesiastical

faith, is still universal in a way no historical faith can ever be. Kant is claiming that no supposed revelation can ever truly command universal *credibility*, even from those who have access to its scriptural sources, since nothing could ever truly justify anyone in being certain that any scripture actually came from God. Rational doctrines, by contrast, to the extent that they can ever be established, are thereby rendered credible universally, for any mind that is capable of considering them and their rational basis. That is the sense in which no scriptural revelation can ever acquire universal authority.

We mentioned in Chapter 1 that Kant was living at a time when critical historical methods of study applied to other historical documents were beginning to be applied to religious scriptures. In this way it was being discovered that the heterogeneity of sources and the history of their composition made religious scriptures at best unreliable witnesses to history – though on political matters they may be as reliable as other sources from the same time. However, the fact that much of popular religion has still not caught up with what was known already in the Enlightenment raises questions about the fundamental intellectual honesty of Christian belief (see Harvey, 1996). Much of the resistance to Kant's *Religion* belongs to the fateful separation in modern culture between popular religion and scientific reason also mentioned in Chapter 1. Kant's project in the *Religion* should be seen as an early, and also a sadly neglected, attempt to heal this open wound from which the culture of modernity still suffers.

The only truly religious function of divine revelation is to provide symbolic representations related to true religion: duties regarded as divine commands. Kant is quite frank about what this involves regarding scriptural interpretation:

This interpretation may often appear to us forced, in view of the text (of the revelation), and be often forced in fact; yet if the text can at all bear it, it must be preferred to a literal interpretation that either contains absolutely nothing for morality, or even works counter to its incentives (R 6:110).

Nor can we charge such interpretations with dishonesty, provided that we do not wish to claim that the meaning we give to the symbols of a popular faith, or even to holy books, is exactly as intended by them, but leave this open and only assume the *possibility* that their authors may be so understood. For the final purpose of even the reading of these holy books, or the investigation of their conduct, is to make better human beings (R 6:111).

Kant is especially clear that no moral authority can be claimed by scripture in direct defiance of what reason tells us is right or virtuous. He cites in this context Psalm 59:11–16 and the moral argument based on it that was made in his time by J. D. Michaelis. These verses celebrate the vengeful triumph over enemies, supposedly with divine assistance, exhorting us to “consume them in wrath, consume them till they are no more” (Ps 59:13).

Kant quotes Michaelis: “The psalms are inspired; if they pray for revenge, then it cannot be wrong: *We should not have a morality holier than the Bible.*” Kant then points out that Michaelis not only offends morality by these

statements but also brings scripture into conflict with itself because Christ commands us to “Love your enemies, bless those who curse you, etc.” (Mt 5:21ff, 44ff).

It is always a danger in ecclesiastical faith that it may not in fact make better people of its members but only harden them in sharing wrong, backward, and pernicious beliefs and vicious character traits. This is a sadly or even an alarmingly common phenomenon among ecclesiastical communities, which we still often see and from which human culture suffers greatly. The only possible response to it is a critical, enlightened (not dogmatic or authoritarian) reception of scriptural documents and ecclesiastical traditions.

In his response to Michaelis, therefore, Kant suggests two alternative readings of the Psalm: one referring to the historical context, and the relation of the Jews to God as sovereign, the other metaphorical, where the intended enemies are not “corporeal enemies, but, symbolized by them, the invisible ones ... evil inclinations, which we must wish to bring under our feet completely” (R 6:110 n). Kant insists that this way of interpreting scripture is nothing new. It is the way all faiths – not only Jewish and Christian but also Greek, Roman, Muslim, and Hindu – have always tended to read their holy books whenever they were read in an enlightened manner:

We shall also find that this is how all types of faith – ancient and new, some written down in holy books – have always been treated, and that rational and thoughtful teachers of the people have kept on interpreting them until, gradually, they brought them, as regards their essential content, into agreement with the universal principles of moral faith (R 6:110–111).

Kant recognizes another authoritative interpreter of scripture besides the practitioner of pure rational (or moral) religion – namely, the “scriptural scholar,” the person in command of the relevant languages and knowledgeable about the history of the people from whom the scripture derives. “*Scriptural scholarship* is required to preserve the authority of a church based on holy scripture, though not that of a religion (for to have universality religion must always be based on reason)” (R 6:112). Scholarship is authoritative, therefore, only *doctrinally* (i.e., for a limited community or church), not authentically for all humanity or for true religion (R 6:114).

Kant firmly rejects, however, the claims of a third pretender to authority in scriptural interpretation – namely, “inner feeling”:

Feeling is private to each individual and cannot be expected of others, even when we have advance cognition of the law from which, or according to which, it arises; thus we cannot extol it as a touchstone for the genuineness of a revelation, since it teaches absolutely nothing, but only contains the manner in which the subject is affected as regards his pleasure and displeasure, and no cognition whatever can be based on this (R 6:114).

It is the function of religion to give symbolic, hence aesthetic and emotional, power to thoughts and principles that matter for human life. But what matters is that the right thoughts and principles should predominate. When feeling –

especially the shared but unenlightened feelings of a religious group – becomes the authority, and sets itself up as superior to our own reason, then true religion turns into its virtual opposite, and becomes an enemy of the human vocation instead of the chief vehicle for its progress, as Kant thinks religion ought to be.

For Kant, therefore, the proper way to interpret scripture is to understand its sense as well as possible in its original linguistic, cultural, and historical context and then seek an interpretation consonant with the spirit of these that is also consonant with the fundamental religious purpose of the act of interpretation – namely, the end of true religion: the self-making of better human beings.

7.7 FAITH AND WORKS: THE ANTINOMY

In the final subsection of division one, Kant takes up the *transition* between ecclesiastical faith and pure religious faith, which has emerged from the earlier discussion as the principal focus of his account of the ethical community and its founding on earth. Existing churches or religious communities are valid – namely, as exemplifications of true religion – only insofar as they harbor within them a process leading from mere statutory service of God, according to the purported empirical revelation of God in a holy scripture, toward pure religious faith grounded on reason. Kant does not immediately take up this transition historically, however – that is left until division two. Instead, he considers the general principle involved in the transition, focusing on one crucial illustration of it, with which he has already dealt in part two, and which we have discussed in Chapters 5 and 6.

As Kant has presented it in the first two parts of the *Religion* (R 6:44–53, 66–84), salvation for a human being beset by a radical propensity to evil depends on a conjunction of two elements: (1) a good moral disposition, grounding a course of improving life-conduct, and (2) faith that God will dismiss the claim against the guilty sinner through vicarious satisfaction or justifying grace. Kant now represents these same two elements as constituting an “antinomy,” based on the precedence of the one element or the other. The question is, Which comes first – the divine motion of atonement toward us, and our faith in it, or our good disposition and resolve to do better? (R 6:116–117). We discussed this topic in Chapter 6 from the standpoint of individual salvation. Kant now raises it in a quite different context: the relation between historical ecclesiastical faith and pure moral faith.

The orthodox Christian doctrine (especially the Lutheran) gives priority to divine grace and our faith in it, arguing that good works must follow from this and that we are incapable of doing anything in the absence of God’s loving assistance. According to orthodox Augustinian theology, to give priority to our efforts is the doctrine of “self-salvation”: the heresy of Pelagianism.

Kant attempts to resolve the antinomy by allowing each side an aspect of the truth. The antinomy as an instance of the tension within true religion between the standpoint of ecclesiastical faith and the standpoint of pure moral religion.

He begins with the insistence that since we can have no cognition of the supersensible, we cannot resolve the antinomy “theoretically,” “through insight into the causal determination of the freedom of the human being.” The matter must instead be considered only “practically,” “whence, in other words, we are to make our start, whether in what God has done for our sake, or from what we ought to do in order to become worthy of it” (R 6:117–118). Here Kant has no hesitation in choosing the second alternative because “no thoughtful person can bring himself to faith [in what God might have done to assist him], however much self-love often transforms into a hope the mere wish for a good, for which one does nothing or can do nothing” (R 6:117). Kant means to avoid the charge of Pelagianism by granting to the first alternative a certain truth as well. He sees the two alternatives, namely, as representing the respective standpoints of historical or ecclesiastical faith and pure religious faith, which must come together in a dynamic of transition from the former to the latter in the faith of any earthly ethical community.

Ecclesiastical faith, being historical, rightly begins with the first principle [of God’s gracious assistance]. But since it contains only the vehicle for the pure faith of religion (in which the true end lies), what in faith (as practical) constitutes the condition, namely the maxim of action, must come first: the maxim of knowledge or theoretical faith must only bring about the consolidation and completion of that maxim of action (R 6:118).

Kant concedes that there is a legitimate standpoint from which faith in divine grace or vicarious satisfaction must come first – namely, the standpoint of ecclesiastical faith, grounded on historical revelation. Kant insists that this must be the starting point of the ethical community as it is found on earth. But the religious legitimacy of ecclesiastical faith itself must be grounded on the progressive process of transition to pure religious faith, from which standpoint a good disposition comes first, and this grounds our hope for divine assistance by making us worthy of it. The priority of faith over works can thus even be seen as a way of representing this transition because ecclesiastical faith is the starting point of the transition and good works are its eventual terminus, even though this lies in the future and may never in fact be fully attained.

Kant argues that in this antinomy (as in the antinomies explored in all three Critiques), the contradiction is only apparent. The two elements needed for saving faith may be seen as merely two ways of representing the same thing. Vicarious satisfaction, namely, is made by the “ideal of humanity well-pleasing to God” (the Christ ideal) which “represents the prototype as the standard measure of our life conduct” (R 6:119, cf. 6:60–66). But Kant has argued that this is really only our way of representing the idea to which our good moral disposition ceaselessly strives to approximate. The two elements of saving faith are therefore “the very same idea, only taken in different relations, as two different principles”; they appear as different only “if one wished to make the historical faith in the actuality of an appearance . . . the condition of the one saving faith” (R 6:119–120).

This, however, would be to ignore the necessary transition of ecclesiastical faith into pure religious faith. In part three, the issue is no longer individual salvation, but rather the historical relation between revealed ecclesiastical faith and pure religious faith. Kant's position concedes to the doctrines of ecclesiastical faith that they may be permissibly accepted as explaining to believers in revealed religion how it is possible for corrupt humanity to become well-pleasing to God through the "foreign satisfaction" (*fremde Genugtuung*) – implicitly: that rendered by Jesus Christ (R 6:117). In that sense, he is willing to allow for the thought that faith must precede works, and therefore leave the antinomy unresolved theoretically. But Kant insists that "practically," when it is a question of "whence we are to make our start," we cannot begin from what God has done on our behalf, but must begin from "what we ought to do to make ourselves worthy of it" (R 6:118).

The misreading of Kant to be resisted here is again presented clearly and forcefully by Wolterstorff:

It seems clear . . . that in Kant's view this is not a true antinomy. Indeed, he himself calls it only "apparent."⁵ For although he lays out both lines of thought . . . he makes it emphatically clear, as he already has in Book II, that in his judgment only the latter is correct . . . As a matter of fact, says Kant, a reflective person cannot bring himself to believe that atonement would be extended to the unconverted; that would be a violation of our moral intuitions . . . God, in the Kantian system, wipes out the guilt of our wrongdoing if we present God with a good character; God is in fact morally required to do so (Wolterstorff, 1991, pp. 50–51).

We saw in Chapter 6 how Wolterstorff's account distorts Kant's views. For Kant, God is *not* "morally required" to find well-pleasing those human beings who have undergone the change of heart. This change makes them morally receptive to God's grace, but Kant does not adopt a conception of justification that reduces God's role to complying with what "our intuitions" tell us is *strictly required* of God. God's forgiving mercy is freely given and not something morally required. God has no duties or obligations because, as sovereign in the realm of ends, his will is not subject to any constraint (G 4:414, 433, 439, cf. KpV 5:32, VpR 28:1066–1068).

What matters for our present concern, however, is the way Wolterstorff's misunderstanding on this point prevents him from seeing what is actually going on in Kant's resolution of the antinomy. It is false to say that Kant rejects the thesis in favor of the antithesis, and this is crucial to the antinomy's resolution. The antinomy arises between what is historically presupposed by the faith of an

⁵ The word "apparent" is an artifact only of the Greene-Hudson translation. What Kant says is, *Hier zeigt sich nun*: "Here [the antinomy] shows itself" (R 6:116). Of course Kant regards all antinomies as *apparent* in the sense that the contradiction can be resolved by seeing how it arises for reason. But this passage offers no support to Wolterstorff's misleading suggestion that this antinomy is any less genuine an antinomy than the others Kant considers in all three Critiques.

actual (Christian) religious community and what this community must do to make itself into a vehicle of true religion, which is its proper vocation.

Wolterstorff's implicit message, as I read it, is that no such reform is necessary. A religion of tradition and authority needs no help from reason. Thus any proposed rational reforms are an alien intrusion to be regarded by Christians only with fear and hostility. Wolterstorff's misunderstanding of Kant's position is needed to make plausible his suggestion that Christian doctrines in a wholly unreformed shape are (as judged, I suppose, by "our moral intuitions") humanly superior to what Wolterstorff thinks Kant is trying to put in place of them.

A more adequate interpretation of Kant would show Kantian rational religion to be more in harmony with historical Christianity than Wolterstorff seems to realize. Kant is not rejecting Christian revelation or trying to put something different in its place. He is advocating instead "the gradual transition of ecclesiastical faith toward the exclusive dominion of pure religious faith" and the latter as a distant harbinger of "the coming of the kingdom of God" (R 6:115). Kant's point is that ecclesiastical faith, without necessarily renouncing its revealed doctrines, ought to seek interpretations of them that bring them into harmony with reason, so that religion becomes the principal way in which people may experience in symbolic terms the moral progress of humanity, which is the proper vocation of the ethical community.

This is why Kant introduces the "antinomy" as presenting the issue "whether a historical (ecclesiastical) faith must always supervene as an essential portion of saving faith over and above the pure religious one, or whether, as a mere vehicle, historical faith will finally pass over, in however distant a future, into pure religious faith" (R 6:116). Those Christians who not only subscribe to the priority of faith over works but insist *practically* on that priority – waiting passively for God's grace and refusing to do anything to be worthy of it – are resisting precisely that reform of their faith that alone can actualize its religious essence. They are – in the terminology Kant will use in part four of the *Religion* – adhering to a "fetish faith" that seeks to conjure up divine action for their own advantage, advocating "counterfeit service" (*Afterdienst*) of God, and transforming the true religion (the inner concentric circle that lies at the heart of their revealed religion) into a priestcraft (*Pfaffentum*) and religious delusion (*Wahn*). What Kant is proposing instead of this is a Christianity that accepts "pure religious faith as the supreme interpreter" of its doctrines and scriptures (R 6:109–114).

Here again we see the falsity of the notion that Kant is rejecting Christianity, attempting to *replace* existing faiths with a pure rational faith even or reduce religion to morality. Kant's position is that we begin with existing ecclesiastical faiths, accepting their alleged revelations and ecclesiastical duties as part of the outer circle of religion. Then, through the interpretation of their symbols, and rational reforms of their teachings and practices, we should seek within them for a true religion of reason and the true ethical community. This parallels

Kant's view that politics should begin with an imperfect or even despotic state and seek to bring it closer to a true condition of right. To depict Kant's project in the *Religion* as an attempt to replace existing ecclesiastical faith with something else is a reaction that would come naturally enough to those who remain stubbornly oblivious to the rational deficiencies in traditional ecclesiastical faith and can perceive no need for reforms in it. Perhaps there are reasons for rejecting Kant's proposed interpretations and reforms or proposing others instead; but these reasons cannot be honestly presented if his fundamental aims are invidiously misunderstood and his actual project is never even addressed.

Kant's position in part three of the *Religion* is that ecclesiastical religion itself can fulfill its proper vocation as true religion only by becoming the empirical vehicle for pure rational religion. The relation between the inner and outer concentric circles of Kant's second experiment is dynamic; this dynamic parallels Kant's conception of human progress in politics and in morality itself. Kant does not reject revealed or ecclesiastical religion but regards it as subject to the vocation of reform in accordance with its true nature as religion.

7.8 THE COMING OF THE KINGDOM OF GOD

Kant ends part three, division one of the *Religion* with a depiction of the terminus of this process of transition, the far off state of affairs in which "the pure faith of religion will rule over all, 'so that God may be all in all' [I *Corinthians* 15:28]" (R 6:121). In part three, division two, Kant sketches a history of Christianity seen as aiming at these reforms. He stresses in this discussion the symbolic nature of religious doctrines and representations, both of God's work and of the human history in which Providence is believed to have shown itself (R 6:134, 136). Kant's project is one he sees as an authentically religious one, even an authentically Christian one. It is malicious misunderstanding to see it in any other light – for instance, as aiming to abolish Christianity and replace it with mere secular morality. Those who can only see it this way should at least admit they are in defiance of Kant's own self-understanding. Kant represents a possible religious future as "a beautiful ideal of the moral world-epoch brought about by the introduction of the true universal religion" and also prophetically "foreseen" (although only symbolically) as the coming Kingdom of God: "'When, therefore, cometh the Kingdom of God?' – The Kingdom of God cometh not in visible form. Neither shall they say: 'Lo here; or lo there!' *For behold, the Kingdom of God is within you*" (Luke 17:21–22) (R 6:136).

The integuments within which the embryo is first formed into a human being must be laid aside if the latter is to see the light of day . . . So long as the human species "was a child, he was clever as a child" and knew how to combine learning too, and even a philosophy helpful to the church, with propositions imposed upon him without any of his doing: "But when he becomes a man, he puts away the childish things" [I *Corinthians* 13:11]. The degrading distinction between laity and clergy ceases, and equality springs from true

freedom, yet without anarchy, for each indeed obeys the law (not the statutory one), which he prescribed himself, yet must regard it at the same time as the will of the world ruler as revealed to him through reason . . . All this is not to be expected from an external revolution . . . – The basis for the transition to a new order of things must lie in the principle of the pure religion of reason, as a revelation (not an empirical one) permanently taking place within all human beings, and this basis, once grasped by mature reflection, will be carried into effect, inasmuch as it is to be a human work, through gradual reform (R 6:121–122, cf. VpR 28:117–119).

Kant's historical prophecies here, and his hopes for religious reform, have often been seen both by the religious and the secular minded as the very reverse of what is either plausible or appealing. To religious people Kant's hopes seem directed toward a watering down of their religion or even its total destruction. They gaze in horror at the end of the "transition" Kant envisions, when there would be no ecclesiastical structure (no clergy) and no religious services ("no statutory observances"). Tradition and scripture, as they see it, would be retained only to the degree that they could be turned into mouthpieces for secular morality. The fundamental doctrines of Christianity (the Incarnation, the Atonement, the primacy of divine grace) would be retained only in the form of strained, metaphorical representations of a secular moral psychology. They might represent metaphorically what Kant would do to their religion by imagining him bulldozing a beautiful cathedral and leaving nothing in its place but an arid moralist's vacant lot. To the secularist, on the other hand, Kant seems hopelessly addicted to outworn metaphors drawn from the Christian religion. Whatever their differences, both sides might agree that nothing like Kant's "transition" from ecclesiastical faith to pure religion has occurred or is likely ever to occur. Moreover, neither side is likely to regard this as regrettable. On that alone they would seem to be in total agreement.

Let me suggest, however, that we consider Kant's prophecies about the future of religion alongside his similar prophecies about the future of the political state. In the latter case, few are likely to judge his views mistaken or condemn his hopes as being for something undesirable. Kant holds that the only civil order consonant with the idea of right is a republican constitution – one involving a separation of powers, a representative legislature, constitutional limits on the executive power, and protection of individual rights (EF 8:349–350, 366, MS 6:311–318, 340). There were few states at the time that exemplified anything resembling this, and Kant himself never lived in one.⁶ But the Kantian position on politics is now widely accepted: that the only truly just state is one whose constitution is that of a representative republic.

⁶ Kant favored a federation among republican states but rejected a world-state, defending a plurality of cultures and ways of life (EF 8:356–357, 367). But in religion he hoped the ethical community would eventually be all embracing, including all human beings as a single people of God. We will return to this difference between Kant's views of the political and the ethical community in the following chapter, when we compare Kant's views with those of Mendelssohn.

Kant's political prophecies may be regarded as astonishingly prescient. But clearly his parallel hopes for religion have not been borne out. Modern culture has tended toward a profound split between religious culture and secular culture. Religion has failed to reform itself; and the mirror image of this failure is that of secular culture's failure to understand and accept what might be beneficial in traditional religion.

Sometimes a philosopher may hold doctrines that are not correct, and the course of history shows them to be wrong. This is true, for example, about many of Kant's moral teachings about sexuality and his views about the social role of women. But I submit that this is not the right way to see the failure of Kant's prophecies regarding religion. For sometimes a philosopher's prophecies may fail not because he is wrong but because the world has gone so terribly wrong. Ecclesiastical religion fails in its *religious* vocation when it becomes an enemy of enlightenment, when it defends traditional, backward ways of thinking rather than leading the way to enlightened social reform. If we do not accept Kant's hopes for the transition of ecclesiastical faith into pure religious faith, we should at least be able to draw from them this latter, more troubling lesson.

Freedom of Conscience

We saw in Chapter 1 that Kant refused on principle to attend religious services at the cathedral where his tomb is now to be found. Part four of the *Religion* is a critique, sometimes even a polemic, directed against many features of traditional religious belief and practice. Kant could never have expected the Prussian censors to give their approval to this part of his book.

Religious services. “Counterfeit service” (*Afterdienst*)¹ of God takes the form of statutory observances, the mechanical reciting of creeds, and formulaic, sycophantic praise of the Deity. Kant rejects as patently self-deceptive the very idea of “means of grace” because although the change of heart can give us a moral receptivity for God’s mercy, grace itself is a free gift from God, which we cannot in principle produce in ourselves or even earn by good conduct (R 6:53, 194). Kant approves of going to church insofar as it is a means of edifying individuals and bringing together the members of a moral community; but he rejects the performance of rituals when they themselves pretend to be genuine service of God (*Gottesdienst*) (R 6:198–199).

Baptism and communion. Kant finds something admirable in activities promoting religious community: welcoming new members, renewing the community’s loving bonds. But baptism, regarded as an “attempt to wash away all sins at once” would be superstitious “thaumaturgy” (R 6:199). He finds admirable the sharing of a common meal as “renewal, under laws of equality” of religious fellowship. But this communion should not be regarded as a “means of grace,” which is a nonsensical and morally offensive concept (R 6:199–200).

¹ “Afterdienst” was a frank German vulgarism used by Luther to criticize the Roman Catholic church of his time. “After” is a word, now fallen into general disuse, meaning “anus.” But Kant uses this insulting word at least as much against Protestants as against Catholics (see R 6:109).

Prayer. We should of course wish for the change of heart in ourselves, and for God's forgiving grace, a wish Kant describes as "the spirit of prayer" (R 6:195). But the external expression of such wishes, as well as of other wishes addressed to the Deity, is at best a superfluous act because God already knows our hearts better than we do. At worst, and all too often, it is a misrepresentation of our inner state, which we are representing to ourselves as more admirable than it really is (R 6:194–197). Petitionary prayer expresses "an absurd and at the same time impudent delusion to have a try whether, through the insistent intrusiveness of our prayer, God might be diverted from the plan of his wisdom (to our present advantage)" (R 6:196 n).²

Creed and catechism. Kant recognizes that creeds belong to existing ecclesiastical faiths but nevertheless raises objections to them. All creeds are fundamentally dishonest because their aim is to pronounce a formula presented to you by others that you must declare you believe but may not even be able to understand. Kant also regards it as inherently unjust for a church to declare a creed unalterable because human beings cannot rationally agree not to enlarge their cognitions or correct their errors (WA 8:390). Kant devises a *moral* catechism; but it is intended to elicit knowledge from the pupil's own understanding (MS 6:480–484). A religious catechism, by contrast, is like a creed: an attempt to inculcate beliefs by imposing them on the catechumen coercively (or at least manipulatively), pressuring a vulnerable pupil to declare them as something he or she believes. About the catechism he was taught in school, Kant remarks, "In childhood we knew it by heart to the last detail and believed we understood it, but the older and more reflective we become, the less we understand of it, and hence we would deserve to be sent back to school, if only we could find someone there (other than ourselves) who understood it better" (Ak 8:323).

Priestcraft. Most of all, Kant denounces "priestcraft" (*Pfaffentum*) – the authority the clergy exercises over the understanding and even the conscience of ordinary religious believers. "I put the main point of enlightenment chiefly in matters of religion because . . . that minority, being the most harmful, is also the most disgraceful of all" (WA 8:41).

This chapter will compare Kant with the Enlightenment thinker I believe most inspired his entire project in the *Religion*. That philosopher was Moses Mendelssohn. We will see that Mendelssohn's practice of his own faith was quite different from what we might expect based on Kant's views as just presented. But Mendelssohn holds on principle that people should not be persuaded to believe things either by threatened harms or by promised benefits,

² When prayer is supplication supported by fawning praise – begging, combined with flattery – it becomes degrading both to God and to us. But it is honest, refreshing, and human to combine self-reproach with accusation of God, as we find in the spoken texts of Leonard Bernstein's *Kaddish* (Symphony no. 3). Prayer then becomes something like a lover's quarrel with God.

whether in this life or another. Kant’s objections to “counterfeit service” can be regarded as his application of similar principles. But the whole matter is much more complex, as we will now see.

8.1 WHAT IS SERVICE OF GOD? WHAT IS “COUNTERFEIT” SERVICE?

From the above it is apparently easy to provide Kant’s answer to these questions. True service of God consists in doing our moral duties, thinking of them as divine commands. It is seeking to effect in ourselves the moral revolution or change of heart, then humbly hoping that God will find our improved life-conduct an atonement that will make us susceptible to his forgiving mercy. Counterfeit service would be anything else people might try to do to become well-pleasing to God, such as all the practices we have just seen Kant denouncing.

That rather simple answer to the questions heading this section would be easy to accept at face value if we were to misread Kant as attempting to replace existing religious beliefs and practices with something else: a religion of pure reason, from which all the objectionable practices of existing faiths have been totally eliminated at one swoop and with dogmatic finality. But we have already seen that this would be a total misunderstanding of his aims. For the religion of pure reason can exist only within a religious community, and no religious community can exist except in the form of an ecclesiastical faith, with its revealed scriptures, creeds, clerical officials, and statutory services of the Deity. This means that Kant’s project must be to advocate the gradual reform of these practices from within. The religion of reason, as Kant presents it, can only point in the direction they should go. There is then an obvious gap between the condition in which churches exist today and the goal of the reform they should undertake. How is that gap to be bridged?

We understand Kant’s project better if we view the religion of reason not as something determinate but only as the sketch or outline of an idea or ideal. It is the pure concept of a single perfected universal ethical community. Therefore, the religion of pure reason cannot yet (perhaps cannot ever) exist in concrete form. More than anything else, this is what makes Kant’s account of the religion of reason – the referent of the title of *Religion within the Boundaries of Mere Reason* – appear impoverished and Kant’s aim falsely appear to be that of “reducing religion to morality.” That false appearance arises because there is as yet little except rational morality in terms of which to describe the aims of the reform of ecclesiastical faith Kant is actually advocating. If pure rational religion is simply the telos of a process of religious reform that Kant hopes will take place in ecclesiastical faiths, then we cannot know what pure rational religion truly is until that process of reform has taken place. Any apparent determinacy of rational religion is therefore misleading. If we let ourselves be misled in this way, it is easy to think that Kant intends us simply to discard the existing forms of ecclesiastical faith and adopt in their place some rationalist

philosopher's idiosyncratic abstraction. That's why it seems Kant is trying to replace religion with something else, not religion at all but instead a secular substitute, of which the common invidious misunderstandings offer us a most unattractive picture.

Once we see the relation between the two circles – the outer one of revealed religion and the inner one of rational religion – in the historically dynamic way Kant intended, it becomes evident that the title of his book refers to a work in progress, the outcome of a process of reform within existing ecclesiastical faith itself, whose outlines are necessarily still hazy. Central to this reform will be an ongoing reinterpretation of the symbols of ecclesiastical faith that Kant hopes will enrich (certainly not impoverish, still less remove or abolish) the religious meanings found in traditional ecclesiastical faith. It will enhance their contribution to our moral lives and to our collective life as an ethical community. For obvious reasons Kant's project focuses on Christianity, especially Lutheran Christianity. But we should suppose that Kant hopes an analogous process to go on in all other ecclesiastical faiths. Religious reform is meant to effect a convergence of traditional religion with modern culture, its science and its morality. This process is to deepen and enrich both traditional religious faith and modern culture.

When we perceive Kant's pure rational religion as an impoverishment of religion, what we are really seeing is only the sad fact that the cultural changes for which Kant was hoping have not occurred. Or if they have, we are refusing to recognize them as something Kant has in mind, perhaps we are also refusing to join in the project of religious reform he is advocating. Our declarations are then not so much descriptions as performatives: we hereby refuse to accept Kant's project, or perhaps hereby refuse even to try to understand it. We are then registering not Kant's failure but instead the failure of modernity, of humanity – that is to say, *our own* failure.³

If we can let ourselves look at things this way, then we can also begin to notice some significant conflicts and uncertainties with which Kant's project in the *Religion* is beset. True religious service of God is *moral* service: but in its religious form, it must symbolize the moral realities of life using the concept of God and related concepts, grasped analogically in a symbolic and emotionally powerful way. This is precisely what distinguishes religion – even, or perhaps especially, the religion of pure reason – from mere secular morality. The symbols are necessarily drawn from the beliefs, practices, and scriptures of existing ecclesiastical faith. These beliefs themselves must already include that

³ We might see Kant's hopes in the nineteenth-century and early twentieth-century Protestant theological movement usually called modernism or liberalism, represented by such figures as Friedrich Daniel Schleiermacher (1768–1834), Albrecht Ritschl (1822–1889), Adolf von Harnack (1851–1930), and Ernst Troeltsch (1865–1923). In 1907 Pope Pius X (later made a saint) officially condemned modernism in the decree *Lamentabili sane exitu*. Today both “modernism” and “liberalism” are swear words for many Christians, both Protestant and Catholic.

symbolism and so must the practices that express those beliefs. In that case we need to ask, Which ecclesiastical doctrines are religious delusions and which are symbolic expressions of true religion? Which ecclesiastical practices, then, which “statutory observances,” really constitute “counterfeit service” of God, and which are ways of presenting symbolically that which encourages true divine service?

Our questions should now no longer have the obvious answers we might have thought they had. Even about “merely statutory observances,” Kant does not hold that they *must* cease; only that they *be able* to cease: “for [they] might always be useful and necessary as a vehicle” of pure moral faith (R 6:135 n, cf. 6:123 n). At this juncture John Hare asks pointedly, “How can historical faith be both necessary and yet able to cease?” (Hare, 1996, p. 46). The official Kantian answer is that they are historically necessary but open to rational interpretations that would make the historical element unnecessary for our religious lives. This answer could not be expected to satisfy those whose hearts still reside entirely in the outer circle of revealed faith and who stubbornly resist any rational reforms that might enhance its religious vocation.

For the inner circle of pure rational religion, however, there is the question, What form will religious symbols and religious practices assume when they become enlightened, including in themselves the understanding and self-interpretation that they are such a vehicle? Some religious practices, perhaps even those about which Kant is most skeptical, might be ways of contributing to our own moral perfection. But of course these same practices can, through self-deception, become nothing but superstitious rituals through which we reinforce pious hypocrisy, subjecting ourselves to an irrational groupthink that perpetuates traditional errors and evils. The difference depends on whether they undergo the reforms for which Kant is hoping.

Analogous questions arise regarding the church as a social institution. As we saw in the last chapter, Kant’s insistence on the vital importance of the moral community to human moral progress shows that he is *not* a moral “individualist.” Kant thinks we need to be part of a community whose members are freely and voluntarily subject to the moral law as a public law and also freely subject themselves to the judgment of other members of the community. Membership in a church by no means contradicts Kant’s claim that we should become “enlightened,” thinking for ourselves rather than falling under the self-incurred tutelage of others. On the contrary, it is the very same thing because it follows from Kant’s own conception of enlightenment that we must look at things from the standpoint of a moral community. Thinking for oneself means thinking according to universal standards, and we can do this only by doing our best to think from the standpoint of all others, becoming “pluralists” rather than “egoists” and using the thoughts of those around us as a *criterium veritatis externum* for our own (O 8:146 n, KU 5:294–295, Anth 7:127–130). Again: individuality and community are not opposed values. On the contrary: Either they are the same value or neither is anything at all.

Today many of us no longer naturally think of the moral community as religious. In Kant's time, however, it was entirely natural to think of it in this way and hard to think of it in any other. There is clearly a great difference between submitting yourself to tribal groupthink in religious matters, becoming part of the swarm that gives *Schwärmerei* its bad name, and being an obstinate contrarian who prides yourself on stubborn idiosyncrasy for its own sake. Obviously we should avoid both these destructive vices and find the mean virtue between them. We should take seriously the beliefs, symbols, and life of our religious community yet also think independently and use our own good judgment about when to defer and when to dissent. This path is no doubt difficult, but that is what Kant is recommending.

It is the same with ecclesiastical authority. Kant opposes in principle the verbal recitation of catechisms and credal formulas and regards the state of self-incurred minority to be at its most degrading when it assumes the form of allowing a pastor or priest to usurp the authority of your own conscience (R 6:189, WA 8:38–39, 41). But this does not entail that the teachings of a church or the advice of clergy must be despised and never allowed to influence your conscientious thinking. The position taken by Kant in the *Religion* is that we ought to be members of a church, freely subjecting ourselves to its doctrines and practices, yet at the same time regarding these critically, in accordance with our own conscience. The task of a church collectively is to view its doctrinal tradition as something to be respected but the symbolism of its teachings and scriptures as objects of constant reflection and reinterpretation. Only by reading the *Religion* in this way can we avoid the error of thinking that Kant intends to do away with existing churches, replacing them with a cult of secular moralism. I hope we now see how that gets Kant wrong. But leaving this error behind raises more questions than it answers.

8.2 MENDELSSOHN AND KANT

The life as well as the thought of Moses Mendelssohn should serve as an inspiration to all who admire the Enlightenment. Mendelssohn was the most prominent figure in the *Haskalah* השכלה (“wisdom,” “erudition,” or Enlightenment). He was born in 1729 to an impoverished family in Dessau. At age fourteen he followed his rabbi, David Fränkel, to Berlin. In Fränkel's yeshiva he learned about the Bible, the Talmud, and the philosophy of Maimonides. The teenage Mendelssohn quickly learned both German and Latin. From this point on, he was largely an autodidact: he was a brilliant student, and in himself he was fortunate to have a brilliant teacher.

Mendelssohn was afflicted with lifelong physical impairments, including curvature of the spine, and also nervous disorders. He was always physically in poor health, and his emotional state was always fragile. It is awe-inspiring, and even a fact hard to comprehend, that despite this, he achieved both business success and intellectual prominence, not only in the Jewish community but far

beyond it. In 1750 Mendelssohn was employed by a successful silk merchant, Isaak Bernhard, to tutor his children and soon became Bernhard's confidant, bookkeeper, and business partner. In 1754 Mendelssohn met Gotthold Ephraim Lessing, and the two became close friends. Mendelssohn was Lessing's model for the title character in *Nathan the Wise* (1779), the great German Enlightenment drama advocating mutual religious toleration among Jews, Christians, and Muslims.

In 1764 Mendelssohn won first prize in a Prussian academy's philosophical-essay competition; second prize went to a then largely unknown Privatdozent from Königsberg named Immanuel Kant. Mendelssohn's generous praise for Kant's essay when the two essays were jointly published, and then Mendelssohn's favorable review of Kant's 1763 essay on the proof for God's existence, constituted the first recognition of Kant as a significant philosopher.

During the 1770s Mendelssohn was drawn into a series of controversies about Judaism and Christianity. The controversy began earlier, with the impudent public challenge issued to Mendelssohn in 1769 by the Swiss religious thinker Johann Kaspar Lavater, declaring that Mendelssohn should either refute Charles Bonnet's defense of Christianity or else abandon the faith of his fathers and convert. This challenge involved Lavater's revelation of personal correspondence that Mendelssohn regarded as confidential. It had a shattering effect on Mendelssohn's mental state and even his physical health.

After Lessing's death in 1781, Friedrich Heinrich Jacobi initiated a controversy with Mendelssohn over Lessing's alleged commitment to Spinozist heterodoxy on religious questions. At first only an exchange of letters, it became public when Jacobi, without Mendelssohn's permission, published their private correspondence in September 1785. The ensuing controversy proved a watershed moment in modern philosophy, bringing Spinoza back into the center of philosophical discussion. But it also brought Mendelssohn into conflict with some of his closest friends, including Herder and Hamann, and took an even heavier toll on his health than the earlier exchange with Lavater. Jacobi is often blamed for Mendelssohn's sudden death in January 1786 at age fifty-six. Kant might have been thought closer to Jacobi's side in the controversy because he rejected Mendelssohn's theoretical proofs for God's existence and immortality of the soul, and like Jacobi, urged faith in them on moral grounds. But Kant's essay of October 1786 (O 8:133–146) offered a sympathetic interpretation of Mendelssohn's rational theology and sternly warned against Jacobi's antirationalist fideism.

We can more easily understand Kant's position in light of the outcome of the earlier exchange between Mendelssohn and Lavater. Together with subsequent publications and controversies concerning the rights of Jews, this culminated in Mendelssohn's treatise *Jerusalem: Or, On Religious Power and Judaism* (1783). *Jerusalem* is divided into two sections: the first deals with the relation of political right to religious freedom; the second expounds Mendelssohn's concept of Judaism, its history and its religious meaning.

Kant's reception of Mendelssohn's *Jerusalem* was enthusiastic. He wrote to Mendelssohn:

How much [do] I admire the penetration, subtlety and wisdom of your *Jerusalem*. I regard this book as the proclamation of a great reform that is slowly impending, a reform that is in store not only for your own people but for other nations as well. You have managed to unite with your religion a degree of freedom of conscience that one would hardly have thought possible and of which no other religion can boast. You have at the same time thoroughly and clearly shown it necessary that every religion have unrestricted freedom of conscience, so that finally even the church will have to consider how to rid itself of everything that burdens and oppresses conscience (Ak 10:347).

This praise of Mendelssohn involves a rather backhanded compliment toward his Judaism because it expresses wonderment that the Jewish faith can be united with so admirable a freedom of conscience. Kant always regarded Judaism as a religion stressing external, morally indifferent statutory observances. Judaism began, he claims, not as a genuine religion at all but only as a political state (R 6:125–126). This picture, however, which some have interpreted as anti-Semitic, was itself derived directly from Mendelssohn, who argues in *Jerusalem* that until the time of King David, there was no distinction in Judaism between political legislation and ethical or religious legislation (1983, pp. 128–130).⁴

Kant's praise of Mendelssohn does, however, contain an implicit thought that Mendelssohn clearly rejected: that the "impending reform" of religion would aim at uniting all churches in a single ecumenical ethical-religious community.⁵ But there is no doubt that Kant approved the version of Judaism Mendelssohn personally practiced and advocated. Part four, second part, §4 of the *Religion* (R 6:185–190) culminates in a vehement defense of freedom of conscience in religious matters, which Kant plainly derives from Mendelssohn's *Jerusalem*. So let us explore it briefly here.

8.3 MENDELSSOHN ON NATURAL RIGHT, CHURCH, AND STATE

Mendelssohn's thought about religion and politics is grounded on a theory of rights and duties that was derived from Pufendorf and the German natural law tradition; according to this theory, we have both natural rights and natural

⁴ Kant holds that the statutory, legalistic phase of Judaism was left behind when Judaism becomes a moral religion (R 6:128, cf. SF 7:53). But there is another aspect of Kant's unflattering picture of legalistic Judaism we should not overlook: he uses it against the Prussian government's attempt to impose Christian orthodoxy by law. This, he suggests (appealing to Mendelssohn), is not true Christianity but only legalistic Judaism (R 6:166 n, SF 7:52 n).

⁵ Kant approved the position of Lazar ben-David, a Jewish Kantian who held that Jews should give up their distinctive way of life in order to be fully accepted into Prussian society (SF 7:54). Mendelssohn of course rejected that idea, which, however, was part of Dohm's advocacy of civil rights for Jews, which Mendelssohn generally welcomed and approved. See §4 below.

duties.⁶ A *right* is a moral authorization to make use of something as a means for promoting one's felicity. The means of promoting our felicity are called *goods*. Our basic perfect rights are to our life, body, and natural capacities. Consequently we have a natural right of property to those *goods* that the use of these capacities creates, or which (as with land and raw materials of labor) are inseparable from the use of our capacities. In a state of nature, independently of any contracts, there is original joint ownership of goods, but there is also a *natural* right of *property* that belongs exclusively to each of us based on the use of our natural capacities (Mendelssohn, 1983, pp. 46–47). What is morally necessary according to the rules of justice is a *duty*. To each of your natural rights corresponds a duty of others. Both rights and duties are divided into perfect and imperfect. Perfect rights are unconditional, and the corresponding duties are coercively enforceable (1983, pp. 45–46). Fulfillment of imperfect duties may not be coerced.

Imperfect rights and duties essentially involve *mutual beneficence*. Mendelssohn maintains that human beings are by nature interdependent. Our needs cannot be satisfied without the beneficence of others; equally, we cannot attain felicity without practicing beneficence toward others. Thus we must both *petition* others for goods and grant their petitions. We petition others for part of their property and grant their petitions by giving them some of our property (1983, pp. 47–49). An imperfect right is a moral capacity to petition another. The one petitioned has an imperfect duty to grant the petition, corresponding to this imperfect right. But granting a petition depends on the free choice of the one petitioned.

Since our capacities are finite, we must devote some of our natural property to our own felicity and can devote only some to the felicity of others. It is up to us how we divide our property between these two uses. Presumably such expressions of human interdependence as market exchanges are a special case of the mutual petitioning and granting of petitions. Imperfect rights and duties can be converted into perfect ones by means of *contracts*. A *contract* is constituted by the reciprocal acts of petitioning and granting a petition (Mendelssohn, 1983, pp. 52–57). If I petition you and you grant my petition, some portion of your goods then becomes my property. My right to it, which prior to this was only imperfect, now becomes my perfect right. It is enforceable against others, including against you. Contracts signal the possibility of a collision between rights: they *create* possible collisions by promoting imperfect rights into perfect ones, which may then come into conflict; but they

⁶ Mendelssohn's conception of perfect and imperfect rights and duties closely follows Pufendorf, 2005, p. 81. Pufendorf himself was relying on a similar distinction already drawn by Grotius. Kant was consciously modifying the traditional distinction (G 4:421 n; cf. Wood, 2017, pp. 46–53). For Kant, all coercible duties are duties of right, and there can be ethical (noncoercible) perfect duties. Kant's distinction has now come to be the one in terms of which we think, so in order to follow Mendelssohn, we must keep our minds open to the older one.

also specify perfect rights and duties so that such collisions may be resolved, as by a court, according to rules of justice.

In addition to contracts involving rights to specific goods, human beings also find it expedient to enter into a *social contract*, by which they create a power capable of enforcing their other contracts and also promoting the common good of society. This power makes possible a more general resolution of collisions between human claims involving perfect rights and duties. The social contract creates a *state*, involving positive legislation, rewards, and punishments aiming at the coercive enforcement of contracts, and the authority to apportion property in a manner that is suitable both to the rights of individuals and the benefit of society (Mendelssohn, 1983, pp. 56–57).

Church and state. Human felicity, as Mendelssohn understands it, is both temporal and eternal. The two are inseparable; the state and the church are equally concerned with both.⁷ (By a “church,” Mendelssohn means any religious society or institution – including the synagogue and the mosque (1983, p. 41).) Church and state equally have the task of educating people and motivating them so that their actions will conform to the demands of duty – both perfect and imperfect duty. Mendelssohn therefore rejects Locke’s idea that the state provides only for temporal felicity, while the church provides only for our eternal felicity (1983, p. 59). The difference instead is that the state is concerned with relations among human beings, while the church is concerned with the human being’s relation to God. It would always be preferable, Mendelssohn thinks, if people did their duty out of conviction rather than through coercion, and both the state and the church try to educate them to this end. But the state may enforce perfect rights and duties, through coercion, while the church should not.

The state gives orders and coerces, religion teaches and persuades. . . . Religious society lays no claim to the *right of coercion*, and cannot obtain it by any possible contract. The *state* possesses *perfect*, the *church* only *imperfect* rights . . . The only aid religion can render to the state is teaching and consoling; that is, in imparting to the citizens, through divine doctrines, such convictions as are conducive to public well-being (1983, p. 45).

Unlike our fellow humans, God needs no service from us. God is not benefited by our praise; God never needs to petition us for any good. The state enforces rights and duties, while the church may neither coerce nor offer benefits as inducements.

God desires no service from us, no sacrifice of our rights for his benefit, no renunciation of our independence for his advantage. His rights never come into conflict with ours. He wants only what is best for us, . . . The words “service”, “honor”, etc. have an entirely

⁷ Mendelssohn thus rejects our U.S. principle of the separation of church and state, but Mendelssohn also holds that there should be no established church: the state should support *all* churches, via taxation of their own professed members. Kant agrees, though he appears to accept an established church, such as the Lutheran church in Prussia (MS 6:327–328).

different meaning when used in reference to God . . . The common man still clings to the accustomed meaning, which has given rise to much confusion in religious matters (1983, p. 59).

For Mendelssohn, as for Kant, it follows that there are no special duties *toward* God. Their expressions of this shared view differ subtly. Kant thinks our duties *regarding* God are really duties to ourselves, to increase our moral perfection by recognizing our human duties as divinely commanded: this *recognition* is Kant's definition of *religion* (MS 6:443–444, R 6:153–154). Mendelssohn holds that we do have duties *toward* God, but because God needs nothing from us, these are not different from our duties to human beings. To the human being who understands that the moral relations of human beings to one another are expressions of God's will, "the moral teachings of reason will be sacred, just as religion is. Nor does religion, or the relation between God and the human being, demand any other duties; it only gives those same duties and obligations a more exalted sanction" (1983, p. 58).

Because God needs nothing from us, and does not petition us, there are not even imperfect rights pertaining to the service or honor of God. Our duties toward God can involve no perfect rights and also no imperfect rights that could ever be converted into perfect ones. Duties toward God can therefore never involve collisions between competing claims of God and of humans. It follows that the church can have no perfect rights whatever. "Where no imperfect rights exist, no collisions of claims arise; and where no decision is required between one set of claims and another, a contract would be an absurdity" (1983, p. 59). Just as Kant was later to do, Mendelssohn contrasts a church with a state mainly through the fact that a state has the rightful power to coerce, whereas membership in a church always remains voluntary. For both philosophers, religious duties depend exclusively on ethical incentives; they are never coercible, either by the state, by the church, or by any power whatever (R 6:93–102).

8.4 MENDELSSOHN'S "NEW AND HARSH" DOCTRINES

All this, Mendelssohn thinks, is self-evident. "All these commonplaces [he says] are so trite that sound common sense wonders how people could ever have been of a different mind. And yet from time immemorial they have acted in opposition to these self-evident principles. Happy will they be if in the year 2240 they cease to act against them" (1983, p. 59).⁸ But an anonymous reviewer of one of Mendelssohn's earlier writings declared these consequences to be "new and harsh." According to this reviewer, Mendelssohn was "denying

⁸ The year 2240 on the Christian calendar corresponds to the year 6000 in Jewish chronology, which also holds that the world will last only six thousand years. So this is Mendelssohn's wry expression of the modest wish that people may come to accept the obvious consequences of these self-evident principles at some time before the end of the world.

first principles” and thus bringing “all dispute to an end” (Mendelssohn, 1983, pp. 199–200). His position is not entitled even to our consideration, much less our agreement. What were Mendelssohn’s “new and harsh” conclusions?⁹

First, churches can make no claims of right. Mendelssohn infers that a church can have no right, *perfect or imperfect*, to goods or property, to contributions, tithes, or renunciations. “The only rights possessed by the church are to admonish, instruct, fortify and comfort” (1983, pp. 59–60). Clergy – ministers, priests, imams, or rabbis – he thinks, cannot be paid wages in the usual sense, but, he argues (following a Jewish tradition) that they can be compensated for their time; this compensation cannot come from the church because it has no rights of property; so it must come from the state (1983, pp. 60–61). Kant would appear to have been influenced by Mendelssohn here, when he holds that among the powers of the state are those of taxing citizens who profess a certain ecclesiastical faith for the purpose of supporting the teaching and divine services performed by that church (MS 6:327–328).

Second, churches can have no rightful or coercive system of government. Mendelssohn asks, “What form of government is therefore advisable for the church?” His answer is simple and blunt: “None!” “For,” he asks rhetorically, “what can be the use of government where there is nothing to govern? What use are authorities where no one is to be a subject? What good is a judiciary where there are no rights or claims to be adjudicated? Neither the state nor the church is authorized to judge in religious matters” (1983, p. 62). The church has only the right to teach and persuade, never to reward or compel (1983, p. 77). Thus there can be no ecclesiastical hierarchy: no cardinals, bishops, presbyters, or councils; no *kahal*. Neither pope nor king nor caliph has religious authority because *there is no such thing as religious authority*. The church has no right either to reward or to punish actions (1983, p. 60), so no ecclesiastical law (1983, pp. 80, 84).

These two “new and harsh” doctrines even brought Mendelssohn into disagreement with Christian Wilhelm von Dohm, military counsellor to Frederick the Great, whose 1781 treatise *On the Civil Advancement of the Jews* Mendelssohn on the whole warmly welcomed. But Mendelssohn disagreed with Dohm on two points: first, Dohm did not favor extending to Jews the rights of political participation; second, he did favor giving Jewish communities the right to expel their members. On the first point, Mendelssohn said, “It will, perhaps, still take centuries of culture and preparation before men understand that privileges on account of religion are neither lawful nor, actually, useful, and

⁹ This anonymous review appeared in *Göttingische Anzeigen von gelehrte Sachen* III (September 1782), pp. 889–894. The piece of writing reviewed was Mendelssohn’s preface to the German translation of Menasseh ben Israel, *Vindiciae Judaeorum* (1656): *Menasseh ben Israel Rettung der Juden, aus dem Englischen übersetzt nebst einer Vorrede von Moses Mendelssohn* (Berlin, 1782).

that it would be a veritable boon totally to abolish all civil discrimination on account of religion" (1983, p. 79). On the second point, Mendelssohn disagreed even more fervently. Dohm argued that any society has the right to expel members. Mendelssohn accepts the principle for societies possessing coercive rights, but where there are no coercive rights, there can be no right of expulsion (1969, pp. 143–144). A church also has no right of expulsion over its members – no excommunication, no Quaker shunning, no *takfir*, no *cherem* such as expelled Spinoza from the Amsterdam synagogue. Of course the *state* can prohibit conduct that interferes with the rightful freedom of citizens to assemble for any voluntarily agreed purposes, including religious purposes. But no church has the right to enforce any rules or laws, or to expel dissenters.

Third, in matters of belief, freedom of conscience is absolute. The state, though not the church, has a coercive right to compel *actions* and omissions that are required to protect the rights of others or provide for the common good of society. When it comes to beliefs and convictions, however, the state and the church are in exactly the same position: neither the state nor the church has the right to use *either punishments or rewards* with the aim of influencing what people believe. When it comes to beliefs or convictions, both church and state have only the right to teach, reason, encourage, and persuade – never the right to bribe or to penalize (1983, pp. 61, 69, 70). Mendelssohn realizes that it may be difficult to distinguish the acts of teaching, encouraging, and reasoning with people from acts of rewarding or penalizing them for agreeing or disagreeing with you. On Mendelssohn's principles, however, this is a crucial distinction – one that the church, as well as the state, must strictly observe. Only reasons, arguments, expressions of feeling, and other methods of teaching and encouraging may be used to influence beliefs and convictions. The use of rewards and punishments, motives of hope and fear, whether in this life or the next, is impermissible when it comes to influencing people's convictions.¹⁰

8.5 FREEDOM OF CONSCIENCE AND EVIDENTIALISM

Mendelssohn's argument for this position might at times seem confused or even self-undermining. He sometimes appears to be arguing that convictions *may not be* compelled because they *cannot be* compelled – because attempts to compel them are necessarily self-defeating. Rewards and punishments, he says – incentives of hope and fear – operate on the *will*, whereas beliefs and

¹⁰ A similar issue arises in Mill with objectionable self-regarding conduct. Others may *express their dislike* of it but must not *resent* or *punish* it (Mill, 1978, p. 93). Kant's views are in substance the same. Kant denies that we may permissibly coerce *ethical* duties or, therefore, punish their violation (MS 6:219–220, 460). We may decline to promote the happiness of others when they count as part of it something of which we disapprove, as long as they do not have a coercible right that we should promote it (MS 6:388). Mendelssohn of course does not expect existing churches to respect these rights, however natural and undeniable they may be.

convictions are the business of the *intellect* or cognitive faculty, which attends only to reasons and evidence, seeks only truth, and shuns only error.

Fear and hope act upon the appetitive urge, rational arguments on the cognitive faculty. You lay hold of the wrong means when you seek to induce men, through fear and hope, to accept or reject certain propositions . . . Whoever wants to act upon the intellect must first lay aside the instruments [of fear and hope]; otherwise he is in danger of proceeding contrary to his own intention (1983, p. 62).

In general, people's convictions pay no heed to benevolence, and are not amenable to any coercion. I cannot renounce any of my convictions out of love for my neighbor; nor can I cede and relinquish to him, out of benevolence, any part of my own power of judgment. I am likewise in no position to arrogate to myself or in any way acquire a right over my neighbor's convictions (1983, p. 61).

A natural response to these arguments is to say that if fear and hope can never influence the intellect, then there is no danger to be avoided, and no wrong to be prevented, by forbidding anyone from attempting to influence beliefs or convictions in such ways. Based on his own arguments, then, Mendelssohn's principles seem redundant. But Mendelssohn's arguments can be better appreciated if we view them in another way. His idea is that the only right and rational way for anyone to form any belief or convictions is to attend solely to the grounds and reasons considered by the intellect. Attempts to influence belief in other ways may or may not succeed, but the very attempt itself aims at the subversion of our rational capacities. That's why every such attempt should be coercively prohibited. "Only the judgment reached by [a person's] powers of intellect can be accepted as valid. To let ideas of good or evil interfere with his deliberations is to permit the matter to be decided by an unauthorized judge" (1983, p. 70).

Here Mendelssohn anticipates Cliffordian evidentialism, which was discussed above in the Preface and also in Chapter 2. According to evidentialism, it is a moral principle that beliefs must follow reasons and evidence; it is wrong to let them be influenced by wishes, hope, or fear. Thus Mendelssohn declares, "I myself can accept as truth only what is rationally convincing to me" (1969, p. 123). We should believe something only because reasons and evidence show it to be true or likely to be true; we should never believe something merely because we think we are *better off* believing it. The evidentialist principle, however, is a principle of private morality. It has nothing to do with our *rights*. Mendelssohn's full position therefore requires a still stronger claim: governing one's beliefs and convictions through the proper exercise of our reason and intellect is also an inalienable right.

Neither church nor state has a right to subject people's principles and convictions to any coercion whatever. Neither church nor state is authorized to connect privileges and rights . . . with principles and convictions, and to weaken through outside influence the power of the truth on the cognitive faculty. Not even the social contract could grant such

a right to either state or church. For to contract concerning things which, by their very nature are inalienable is intrinsically invalid, and cancels itself (1983, p. 70).

This right is what Kant has in mind when he praises Mendelssohn's defense of "freedom of conscience." It also provides us with a different way of understanding Mendelssohn's arguments that the use of reward and punishment, hope and fear, to influence beliefs is necessarily ineffective or even counterproductive. His claim is that if people's convictions were formed as they should be, according to reasons and evidence, then attempts to influence them in these ways *would always be* ineffective or even counterproductive. Motives of hope and fear can influence a person's beliefs and convictions only by violating the person's inalienable right to form convictions free from such wrongful influences. Beliefs should not be based on feared penalty or hoped for advantage, whether it is monetary (for believing seven is not a prime number), eternal (heavenly bliss), or moral (increased moral motivation).

Churches and states alike tend to violate these inalienable human rights by rewarding or punishing the *expression* of beliefs and convictions, as well as by prohibiting or interfering with the critical examination of the beliefs they seek to impose. One might therefore think that freedom of conscience is adequately protected by protecting freedom of *speech*. This thought is mistaken, however, for at least two reasons. First, freedom of speech is usually thought to be violated only through the coercive *prevention* or *prohibition* of the expression of beliefs, not by offering rewards or inducements. Mendelssohn's position would prohibit the use of rewards as well as penalties and for exactly the same reasons. Second, Mendelssohn, like Kant, realizes that the public good may require some limitations on free expression. Both philosophers would doubtless approve of some state limitations on free expression that we probably would not. Mendelssohn thinks that the state may *discourage* the spread of religious doctrines – specifically, "atheism and fanaticism (*Schwärmerei*)" – that are judged contrary to the good of society. But he insists that it should do so only noncoercively, "only from a distance" and with "moderation": "It should not interfere in any dispute or wish to decide it through the use of its authority" (1983, p. 63). Further: "There is not a single point in the entire sum of human knowledge that is to be placed beyond question, not one iota that may not be subjected to investigation" (1983, p. 81).

Mendelssohn's point is that limitations on free expression must *never* be put in place with either the aim or the effect of undermining people's formation of beliefs purely on intellectual grounds. Neither the letter of the U.S. Constitution's First Amendment nor the spirit of its application protects freedom of conscience in this sense. We make no distinction between talk that promotes reasoned conviction and talk that makes it harder to achieve. Often in American society, economic factors (and advertising) favor the dissemination of unreasoned convictions and propaganda (communications that undermine the values supporting rational conviction). Admittedly, this is a hard distinction to

draw in practice, and trying to enforce it coercively would be directly destructive of freedom of conscience. There is still an important difference in practice between honoring the spirit of freedom of conscience and treating it with contempt.

8.6 MENDELSSOHN'S ENLIGHTENED JUDAISM

My theme here – in this book, even in this chapter, is not the meaning of Judaism; on the contrary, it is mainly Kant's understanding of Christianity. But since the interpretation and defense of Judaism meant so much to Mendelssohn and since his conception of Jewish religion sheds light on his conception of freedom of religious conscience, I will briefly expound his understanding of Judaism, as I have been trying to do with Kant's understanding of Christianity. The entire second section of *Jerusalem* is devoted principally to articulating Mendelssohn's conception of the Jewish religion. Mendelssohn had been accused of adopting principles regarding religion and the state that are inconsistent with the faith of his fathers and might even require him, in the spirit of Lavater's earlier challenge, to convert to Christianity. His deepest aim in *Jerusalem*, Mendelssohn says, is to answer these charges (1983, p. 84).¹¹

Mendelssohn maintains that Judaism, unlike Christianity, has no revealed doctrines, no articles of faith, nothing that one must *believe* in order to be saved. "Judaism knows of no revealed religion in the sense in which Christians understand the term" (1983, pp. 89–90). The beliefs on which Judaism rests are simply those eternal truths that can be arrived at by the human intellect after free and honest examination. They are knowable by reason alone (1983, p. 94, 1969, p. 128). The existence of God, divine providence, the immortality of the soul are in Mendelssohn's view, "not merely comprehensible to human reason but can also be demonstrated and verified by human powers" (1983, p. 89, cf. 1969, pp. 128–129, 148). For the Jew, "there is no conflict between religion and reason, no rebellion of our natural cognitive faculty against the stifling authority of faith" (1969, pp. 154–155).

Judaism has no revealed *doctrines*, but of course it does have a revealed *law* – Mosaic law – written in the Torah and interpreted through the ages by oral and written traditions and practices (Mendelssohn, 1969, p. 97). This law prescribes only deeds, never beliefs. "Among the prescriptions of the Mosaic Law, there is not a single one which says: *You shall believe or not believe*. They all say: *You*

¹¹ The specific challenges to Mendelssohn subsequent to Lavater's were those issued by the anonymous author of *Das Forschen nach Licht und Recht in einem Schreiben an Herrn Moses Mendelssohn auf Veranlassung seiner merkwürdigen Vorrede zu Manasseh von Israel* (Berlin, 1782) and a postscript to it by Daniel Ernst Mörschel. The author of the brochure was at the time reputed to be the Austrian jurist Joseph von Sonnenfels, but was probably August Friedrich Cranz. In *Jerusalem*, Mendelssohn refers to the author simply as the "Researcher" (*Forscher*).

shall do or not do. Faith is not commanded, for it accepts no other commands than those that come to it by way of conviction" (1983, p. 100). Mendelssohn was a famously strict observer of Jewish laws and traditions. But he does not think that Jews may be *coerced* to follow them or *rewarded* for following them, and he does not regard them as binding in any way at all on non-Jews. "Anyone not born into our community need not observe its laws" (1969, p. 117). More generally, Mendelssohn holds that "no particular religion can be exclusively true" (1969, p. 117).¹²

The purpose of Mosaic law, according to Mendelssohn, was to insure the felicity of a particular people at a particular time, including their descendants. God chose this people to be a witness to his righteousness and power; its aim is only the felicity of the Jewish people.¹³ This law and these traditions are the foundation of national cohesion and community (1983, pp. 93, 99, 127). "It is the best religion for ourselves and our descendants, the best for certain times, circumstances and conditions" (1969, p. 134). "We believe the laws of Moses are strictly binding upon us as long as God does not revoke them explicitly . . . Divine laws remain unalterable until we can be utterly sure that God himself has announced a change" (1969, p. 155).

Not all Jews, either in Mendelssohn's day or in ours, would accept his version of Judaism. Some have charged that it was consciously inconsistent, even disingenuous (e.g., see Arkush, 1994). I see no reason to doubt that Mendelssohn was entirely sincere in arguing that the religion of his fathers, if rationally interpreted, is completely consistent with rational principles and religious freedom. But as we have seen, analogous charges of guile and deception have been lodged against Kant. I would even go so far as to suggest that Kant's project in the *Religion* may even have been suggested to him by Mendelssohn. Kant's *Religion* attempts to do for Christianity what he thought Mendelssohn had succeeded in doing for Judaism.

8.7 WHERE MENDELSSOHN AND KANT DIFFER

Mendelssohn's evidentialism is much simpler than Kant's, and even more radical than Clifford's, since Mendelssohn maintains that there is not only an ethical duty to believe only according to reason and evidence but even a right – in principle,

¹² Kant would put it this way: there is one true religion (namely, recognizing your duties as divine commands), but no particular ecclesiastical faith can be exclusively true.

¹³ Mendelssohn does seem committed to regarding the witness of scripture and tradition as providing sufficient theoretical evidence for the claim that the Mosaic law has been revealed by God. Even supposing that Mendelssohn is right that we can prove God's existence, it seems highly questionable whether this can be justified. The works of Reimarus, published through the efforts of Mendelssohn's friend Lessing, certainly call into question the reliability of the texts and traditions to which Mendelssohn must appeal. Precisely on these grounds Kant avoids appealing to scripture or ecclesiastical tradition for historical claims, insisting that such reports be interpreted symbolically for moral purposes, not relied upon for history or natural science.

even a coercively enforceable right – not to have others attempt either to coerce you to believe or not believe something or even to offer you inducements (whether temporal or eternal) to believe or not believe. If “faith” means *belief*, then it would be wrong to tell people that they can be saved by faith. Mendelssohn accommodates this position by regarding the existence of God and immortality of the soul as theoretically demonstrable by the very arguments that Kant seeks to undermine in the *Critique of Pure Reason* (Mendelssohn, 2011, 2006). It would require major changes in Mendelssohn’s philosophy to abandon these theoretical arguments.¹⁴

It was, and remains, controversial for Mendelssohn to say there are no revealed doctrines in Judaism. But it would have been merely absurd for Kant to declare that there are no revealed doctrines in Christianity; and of course he makes no attempt to do this. As we have seen, Kant’s position regarding revelation is necessarily subtle and easily misunderstood. No more than Mendelssohn does Kant regard revealed doctrines as theoretically cognizable by human reason. When Christians so regard them, Kant must treat their beliefs as theoretically unsupported, religiously permissible as part of the outer circle of revealed faith, but not required for a saving faith, and best understood religiously as symbolic expressions of aspects of the moral life as it can be experienced by pure rational religion. Kant’s version of evidentialism must therefore be less simple, allowing for Belief (rational assent on practical grounds), distinct from belief in the ordinary sense, habitual unconditional assent according to theoretical evidence (which we discussed in Chapter 2, §§ 8–9). For the same reason, it is too easy, as we have seen, to misunderstand Kant’s position as outright rejection of important parts of Christian doctrine.

The relation of religion to ethics must also differ between the two philosophers. Mendelssohn was a Wolffian perfectionist or eudaemonist in ethics, holding that the morally right course of action is also that which promotes the agent’s happiness. Mendelssohn’s view of the Jewish law he observed with famous strictness was that it was God’s way of insuring the morality and also the happiness of a certain people at a certain time and also the community of their descendants. Kant may not have accepted this specific claim, but he would have respected the right of an individual or a people both to apply the moral law according to its own judgment and to seek its happiness according to its own lights; and it would have been perfectly consistent with Kantian ethics for individuals, or even an ecclesiastical community, to judge that the correct application of moral principles consists in following certain traditional customs. Kant would condemn this as *Afterdienst* only if statutory

¹⁴ Mendelssohn famously describes Kant’s critique of these proofs as “world-crushing” (*weltzer-malmend*). Given Mendelssohn’s position, it would have been world-crushing *for him* to accept Kant’s critique of these proofs. This famous remark should always be read as above all an admirable expression of Mendelssohn’s characteristic frankness and open-hearted honesty. No wonder he responded with the *Morgenstunden*, a defense of his rational belief in God.

observances are regarded as morally binding independently of rational morality; but Mendelssohn would not have represented the Mosaic law in that way. Obeying the law of Moses is the way for a Jew to be both virtuous and happy.

Kant's hopes for future religious reform included the eventual unity of ecclesiastical faiths in a single ethical community. The relation of this doctrine to the traditional Christian hope for the conversion of all humanity (including the Jews) is evident here, and Kant wants to use this point in his appeal to his Christian audience. Mendelssohn, however, defended the permissibility of Jewish separateness. Mendelssohn also sees no reason why there should be a "union of faiths," as suggested by Kant's declaration that "humanity will finally be united with regard to the essential point of religion" (Ak 10:347). But here the two philosophers need not ultimately disagree. Kant says there is only one true religion, though there may be many ecclesiastical faiths which serve as its vehicles (R 6:107–108). The differences among faiths – their differing scriptures, revealed doctrines, or laws, their differing statutory practices, might continue to exist and serve as differing vehicles for pure moral religion. For recall that Kant's claim is not that historical faith *must actually* cease, only that it *be able* to cease, in the sense that religious people come to realize that their purpose is always "only the intrinsic firmness of pure moral faith" (R 6:135 n). It would be consistent with this that different faiths might evolve in such a way that they can "be united with regard to the essential point of religion" yet without abandoning their distinctive traditions. That's a hope that Mendelssohn and Kant could share.¹⁵

It was no illusion on Kant's part when he praised Mendelssohn's *Jerusalem* for defending the principle of "freedom of conscience." What most impressed Kant in Mendelssohn's position, I think, is its most extreme evidentialist claim, which makes belief or assent according to reason alone a matter of right: neither civil nor ecclesiastical authorities should be permitted either to offer rewards or to impose coercive sanctions on individuals regarding the content of their beliefs. That is what I think Kant meant by the "freedom of conscience" defended by Mendelssohn, of which he said that that "one would hardly have thought it possible [to unite it with Judaism] and of which no other religion can boast" (Ak 10:347). This last clause ceases to single out Judaism and asserts simply that Mendelssohn's principles would be difficult to unite with *any* ecclesiastical faith. Even an enlightened Christianity, Kant is suggesting, would be unable to boast of them. The Kantian version of Mendelssohn's idea, however, is expressed in terms of *conscience* – an important theme in Kant's ethics even independently of his views on religion.

¹⁵ Yet a third alternative, still in essential harmony with the spirit of both Kant and Mendelssohn, was later proposed by Franz Rosenzweig. He sees different religious traditions, in particular the Jewish and the Christian, as bearing the same witness to the truth yet in different and complementary ways. He thinks their differences will fall away in history as they achieve the ultimate redemption of the world (Rosenzweig, 1985, pts. II and III, pp. 205–424).

8.8 KANT ON RELIGIOUS CONSCIENCE

Kant's ethical theory is based on a single principle or law, which he formulates in a variety of ways (see G 4:436). None of them, however, offer us a moral decision procedure or a direct, discursive criterion of right and wrong action.¹⁶ The principle grounds a plurality of moral duties or concepts of virtue (and vice). It is the task of judgment to apply these concepts correctly to particular cases, by deciding which actions are permitted, forbidden, and required by them. Kant's definitive formula of the moral principle is the formula of autonomy and its more intuitive variant, the formula of the realm of ends (G 4:431–434). Particular duties are derived from the formula of humanity as end in itself (G 4:429–431). The first and most often discussed formula, the formula of universal law, and its variant, the formula of the law of nature, have the modest function of guiding judgment in applying duties and especially in serving as a standard or canon that can (sometimes) save us from mistaken applications based on our human tendency to self-preference and self-indulgence (which are not difficult to recognize as aspects of the radical propensity to evil) (G 4:403–405, 421–424).

Judgment is a talent, partly inborn and partly developed through experience. It is not reducible to any procedure of discursive reasoning. We either apply a concept correctly or we apply it incorrectly: that's all there is to it (KrV A132–134/B171–174; TP 8:275, Anth 7:199). Moral judgment, in Kant's view, is a task of the understanding. It is fallible, not only because it may be corrupted by self-preference but also because our capacity for correct judgment is limited. It is our task as moral agents to do the best we can to avoid errors of moral judgment – not only to choose actions that are not objectively wrong and not to omit actions that are obligatory but also to choose actions that both display and develop moral virtues, while avoiding vices. But if you are a wise moral agent, you know that your best is not always good enough. It is not always possible for the understanding to judge correctly about right and wrong. Especially in cases where hard moral decisions are to be made, you know your judgment is fallible and your conclusions are uncertain. But you do know that you must act even in conditions of moral uncertainty. That's where conscience comes in.¹⁷ In the *Religion*, Kant puts the resulting moral predicament this way:

Now it is the understanding, not conscience, which decides whether an action is in general right or wrong. And it is not absolutely necessary to know, of all possible actions whether they are right or wrong. With respect to the action that I want to undertake, however, I must not only judge, and be of the opinion, that it is right. I must also be *certain* that it is (R 6:186).

¹⁶ For my defense of this last claim, and of most of the things I am about to say in the next few paragraphs, see Wood (2017).

¹⁷ For more on conscience and judgment, see Wood (2008), especially chaps. 3 and 10, and again, Wood (2017), pp. 13–60. My thinking about Kant on conscience was much enriched by discussions I had on this subject with Samuel Kahn, especially during the years 2009–12, when he was completing his Stanford PhD dissertation with me, partly on this topic.

I cannot be *intellectually* certain that my action is objectively right, but I can (and must) do my best to avoid error and especially to avoid self-deception about my own maxims and disposition. If I am to act, and act with certainty, I must judge that I have done my best, even where I know my best is not always good enough to insure my doing the objectively right thing. A second judgment, therefore, is required.

A judgment of my understanding, which is fallible even when I have done my best, identifies which action I ought to do and which to avoid. If I am to act, I must make a second judgment: I must judge *with certainty* that I have done my best to determine what is right and wrong and that I have not let my intellectual deficiencies, my vices, or my self-preference influence this first judgment, at least to any extent beyond that which I absolutely cannot avoid. To decide what is right, I pass a judgment of the understanding. This judgment is always fallible and may be innocently mistaken. This fallibility, however, itself constitutes for me a serious moral problem. For to act with moral resoluteness and certainty, I cannot permit my fallibility to offer me an excuse for moral frivolity or take my duties less seriously on its account. This means I must pass a second judgment on myself in making this first judgment – I must judge with certainty that I have done my best to determine what is right and not to let that judgment be corrupted.

This second judgment is what Kant means by the judgment of *conscience*.

Conscience could also be defined as *the moral faculty of judgment passing judgment on itself*, except that this definition would be much in need of prior clarification of the concepts contained in it. Conscience does not pass judgment on actions as cases that stand under the law, for this reason does . . . Rather, here reason judges itself, whether it has actually undertaken, with diligence, that examination of actions (whether they are right or wrong), and it calls upon the human being himself to witness for or against himself whether this has taken place or not (R 6:186).

In the *Metaphysics of Morals*, Kant describes conscience as an inner court, in which the moral agent plays all the roles: accused, prosecutor, defender, witness, and judge (MS 6:400–401, 437–442, VE Collins 27:351–357, VE Vigilantius 27:613–620). If the process has been fair and honest, conscience is infallible in the sense that there is no moral appeal beyond it: it is the moral court of last instance. Like any other court of last instance, we may later decide that its verdict was objectively wrong, but we cannot appeal beyond it regarding the individual action. In the case of conscience, we may later decide that the action we did in good conscience was objectively wrong; but we must not say that at the time we ought to have acted against our conscience.¹⁸

Kant thinks we cannot fail to pass this judgment on ourselves because to impute an action to oneself is the most basic function of conscience (VE

¹⁸ The U.S. Supreme Court has obviously made many objectively wrong decisions, and with recent appointments, we must dread still more of them to come. But each decision is forever legally valid for *that* case, even if its import is overturned for future cases by legislation or by later decisions. Likewise, Kant holds that any conscientious choice must be judged right for the agent at the time, even if others, or the agent him- or herself later on, decides it was made in error.

Vigilantius 27:575). Conscience is in this respect like an instinct, which operates in us whether we will or not (VE Collins 27:297, 351). If I literally had no conscience, if I did not pass such a verdict on my actions, then I could not impute an action to myself and would not be a moral agent at all. But I can (and people often do) fail to heed the verdict of conscience and deceive myself into thinking it has rendered one verdict (e.g., my innocence) when in fact it has rendered another (my guilt). This is what he thinks we mean when we say that someone is unconscientious or lacks a conscience. When the verdict is not heeded, or a dishonest verdict substituted in its place, conscience has not erred because what we are talking about is not a genuine verdict of conscience (MS 6:401). If you have pronounced a genuine verdict of conscience, and followed it, Kant holds that nothing more can be morally required of you, even if your judgment about the objective rightness of an action has been in error (MS 6:401).

Kant takes seriously the traditional idea that the voice of conscience is the voice of God. He denies that conscience gives us any ground for assuming the existence of a Deity; but if I am religious, and therefore think of my duties as commanded by God, I can “by analogy” represent the judgment of conscience as God’s judgment (MS 6:440, VE Collins 27:296–298, VE Vigilantius 27:574–575). Religious thinking is treated as symbolism, using analogy – that is, we think about our own most accurate and fair self-judgment in the same way we think of God judging us. We present (*darstellen*) our accountability to ourselves as our accountability to God.

When conscience is understood in this way, it becomes evident why Kant believes in absolute freedom of conscience. He agrees with Mendelssohn that in religious matters, no authority, civil or ecclesiastical, has the right to offer external incentives (rewards or penalties) for believing or withholding belief. Believing or assenting to a religious proposition is not an external action that could violate the right and be subject to coercive penalties; and there is no way that any external authority could be in a position to judge whether the person’s belief or lack of belief accords with their own judgment of conscience (R 6:186). A conscientious belief, whether correct or incorrect, could never be an object even of moral blame, still less a crime deserving of punishment; and to reward someone for violating their own conscience is always wrong. Even if we assume their belief is objectively wrong, no external incentive should be offered to change it: “It is certain that to take a human being’s life because of his religious faith is wrong,” unless the inquisitor could know somehow empirically that God has willed otherwise. “But that God has ever manifested his awful will is a matter of historical documentation and never apodictically certain” (R 6:187).

Kant ridicules the argument that a government cannot violate conscience by requiring or prohibiting public declaration of religious beliefs on the ground that it cannot thereby prevent people from thinking whatever they please. “We make fun of this, saying that no freedom is thereby granted by the government, ... [yet] what the secular power cannot do, the spiritual power

can. It can prohibit even thought, and actually hinder it as well; indeed, it can exercise this coercion (namely, the prohibition even to think otherwise that it prescribes) upon its mighty [ecclesiastical] authorities themselves.” It does so by encouraging in human beings “the propensity to a servile faith” in which people are taught that God punishes them even for thinking freely. “True, to be free of this coercion one needs only to will; . . . but it is precisely this willing on which a bar is being placed internally” (R 6:133–134 n, cf. O 8:145). Here the sometimes sweet and touching thought that there is an irrationality above reason loses whatever charm it might ever have possessed and at last shows itself unmasked in all its cruel and loathsome obscenity.

The first step in reforming ecclesiastical faith must therefore be that of encouraging individuals to follow their conscience in religious matters. Kant thinks that “internal coercion will disappear on its own” once people are released from the external coercion of institutional authority, “which hinders all spontaneous advances in the ethical communion of the believers.” It is a religious community characterized by this freedom of conscience that “constitutes the essence of the true church” (R 6:133–134 n).

As this shows, individuals themselves violate their own freedom of conscience whenever they submit to a servile faith – a faith that accepts promises or threats as reasons for believing or not believing. Even if we grant that there are revealed laws or doctrines, it is wrong either to require or reward people for their public profession; this is especially true in the case of mysteries, where someone is given or accepts an external incentive for professing what they know they cannot even understand (R 6:187). Particularly abhorrent is what Kant calls the “maxim of security” (*argumentum a tuto*), which reasons that if I will be rewarded for professing some belief, then I should do so even if it is not approved by my conscience because I then risk nothing in order to obtain the reward, and the worst that can happen is that the belief is false. “*The violation of conscience*,” Kant says, “in proclaiming as certain, even before God, something of which he is yet conscious that, its nature being what it is, cannot be asserted with unconditional confidence – this *the hypocrite regards as a mere nothing*” (R 6:188–189).

Let us hope the time is past when people are burnt at the stake for professing what they believe. But it is still true that people are offered the inducements of approval and fellowship for professing whatever beliefs their tribal community approves. This is often a mechanism through which beliefs held by a community that are lacking evidential support, or even held in the very teeth of the evidence, are reinforced among its members. This seems to be true, moreover, of political as well as religious beliefs, and although it is far more common among some communities and for some beliefs than it is of others, it is a violation of freedom of conscience that still prevails widely and even receives general acceptance.

Here again, however, we must not make the mistake of thinking that conscience would necessarily reinforce “individualism” in the sense of stubborn idiosyncrasy, at the expense of community, or require rejection of any and all beliefs that tie a community together. A person’s honest judgment may be, and ought to be, influenced by their community and by the authority they may reasonably, though always critically, grant it over their opinions. A conscientious person will weigh the authority of this source along with evidence and reasons coming from other sources and form their beliefs in light of all relevant considerations. It is all too easy, as we have seen, to pose the impossible dilemma of, on the one hand, submitting uncritically to the tutelage of one’s community and, on the other, refusing on principle to be part of a community of faith. It is then equally easy to impose this dilemma on Kant and conclude that the religion of reason and the principle of freedom of conscience must be incompatible with institutional religion in general.

Katrin Flikschuh makes this mistake when she argues, “Conscientious but differing and inconsistent private judgments are possible. Then we follow the judgment held by the community – or its individual members; but if each follows his own judgment, then there is no longer a community” (2010, p. 210). This is true, of course, and even unavoidable, when communities hold the doctrines that unite them not only contrary to reason but also dogmatically and intolerantly, in the manner of an unenlightened and servile faith. Both Kant and Mendelssohn, as advocates of Enlightenment, are hoping that churches will reform themselves in ways that respect the freedom of conscience of their members. A religious community that respects the freedom of conscience of its members will neither reward them for conforming to its beliefs nor punish them for failing to do so. It will welcome reason, not reject it. It will adopt only reasonable beliefs and offer reasonable interpretations of its symbols, on which its members may freely reach consensus based solely on their reasonableness. Such a church will also be open to diversity, and its members will, both within their own minds and consciences and also in their interactions with one another, encourage questioning, revising and improving both their own beliefs and those they share with others that constitute their community of faith. The views of Mendelssohn and Kant about freedom of conscience in a religious community are being addressed to churches that accept these principles or at least are on the way to accepting them. In such a church, it will be precisely through all members following their own conscientious judgment that the existence of the religious community is perpetuated.

We don’t see many religious communities of this kind around us, though, do we? (Or do we? I don’t pretend to judge all religious communities, especially those I don’t know. Perhaps some religious communities do approach this desirable condition.) But if we think of religious communities as inherently opposed to enlightenment and freedom of thought, then we naturally think Kant’s advocacy of freedom of conscience is incompatible with his insistence that religion must take the form of an institution or a community. We will be

even more likely to do this if we think there is some ultimate incompatibility between fulfilling our individuality and belonging to any human community. But that thought is moral despair. The Kantian view is that individuality is fulfilled only in community, and a true community must be a community of free, self-governing individuals.

Here again, Kant's teachings in the *Religion* will appear hopelessly indefensible as long as we tacitly assume the falsity of the basic principles on which they are grounded and complacently take for granted the failure of all the historical hopes they cherished.

Concluding Remarks

This book has tried to present Kant's project in the *Religion* as he saw it and to offer a sympathetic but not uncritical exposition of its chief claims and arguments. Kant had an ambivalent relationship with his own religious upbringing and the religious practices of his own time and place. But we misunderstand his attitude if we take it to be a rejection of religion. Kant most certainly did not intend to embrace "secularism" as opposed to religion.

Kant saw his moral argument for God's existence as an important part of his philosophy. In my presentation of that argument, I have tried to stress the soundness of the basic idea behind it – namely, the existence of distinctively practical grounds for assent to that on which we base the hopes for our ultimate concerns (what Josiah Royce called our "loyalties"). I've argued that the sound basis of the argument is something even secularists ought to acknowledge, regarding their own social and historical commitments. The moral argument can therefore also constitute, as it did for Kant, a serious and substantial basis for commitments and hopes that take a religious form. It is by no means a half-hearted substitute for the theistic proofs that convince few today and that Kant famously criticized. People badly underestimate Kant's moral argument because they do not understand or properly appreciate its sound rational basis in the rational relation between purposive action and theoretical assent to its presuppositions.

Kant's philosophy grounds morality on human reason, not on religious belief. Enlightenment modernity cannot take morality seriously, or even correctly conceive of it, as long as morality is combined with the slavish and authoritarian attitude that morality must collapse unless it proceeds from the arbitrary will of a cosmic tyrant. Kant argues, however, that rational morality is not only compatible with belief in God but even leads to the recognition of our moral duties as divinely commanded. This recognition, which Kant calls

“religion,” is distinct from morality. Kant holds that religion is supported by morality but goes beyond it.

In *Religion within the Boundaries of Mere Reason*, Kant distinguishes two kinds of religion: a religion of pure reason and a religion based on divine revelation. Revealed religion is an outer concentric circle, which can receive rational support from the religion of pure reason, conceived as the inner circle. The two circles are even mutually dependent because for Kant the religion of the inner circle depends historically for its existence on moral communities of faith, hence on the revealed religion of the outer circle, while the inner circle constitutes what is most essentially religious about the outer circle. Kant holds that both circles of faith will be enriched, and the distinction between them even blurred, by a historical process of reform in which revealed religion comes to appreciate its rational core and comes to interpret its own beliefs and practices in a way that harmonizes with reason, science, and enlightenment. Rational morality will understand itself better if it sees itself in relation to its historical basis in religion – not only in its distance from that history in religious superstition but also in the way that its historical religious basis can deepen and enrich the moral life.

All religious thinking is necessarily *symbolic*. This holds at least as much for religion that remains oblivious to its essentially symbolic character as for religion that acknowledges and welcomes it. For it is only symbolic thought that makes religious beliefs and emotions meaningful to people’s lives. Kant holds that the revealed doctrines of the outer circle of religion can be seen as symbolizing, and thereby enhancing, the symbolic presentations found in pure rational religion. Revealed religion, however, as symbolism must be distinguished from the literal truths of science, history, and morality. Once it is so distinguished, the literal truths and the symbolic religious truths can be seen to be compatible and also harmonious, fulfilling different human needs and answering to different rational requirements. If Kant errs at this point, it is in his overestimation of the capacity of religious thought and institutions to develop and reform. But that is more a failure of religious thought and institutions than it is Kant’s error. We have not lived up to Kant’s hopes.

Kant’s project in the *Religion* is based on the conviction that the vocation of all religion is to make better human beings of us by promoting goodness of heart (love and respect toward others) and the struggle against our temptations to evil. Human progress rests on human community and leads in the direction of a universal moral community. Unlike most cosmopolitans since his day, Kant does not see the human community as a *secular* community as opposed to a religious one. He sees the universal moral community as a “people of God,” even a people fulfilling the Christian hope for a church triumphant or the coming of a Kingdom of God on earth. In Kant’s view, existing churches, whether Christian or non-Christian, can join in this community, at least if their self-interpretation undergoes the right rational reform.

One place where Kant may go wrong is in his assumption that existing differences in faith and ethnicity may be overcome through absorption in a universal community. The chief danger is that a majority or dominant community, such as the Christian one in Kant's time and place, may impose its ways and its symbols on others. We should therefore take seriously Mendelssohn's defense of an independent way of life for different ecclesiastical communities. To put it in the context of Kant's time (which is still not so different from ours), first, the majority Protestants (or Lutherans) must learn to respect Jews, Muslims, and Catholics as human beings, as individuals with their own traditions and ethnic identities. Only then can there be any question of a universal or ecumenical moral community. The same would apply in our time and place to respecting all peoples of color first as who they are here and now. Only then may we presume to hope for a community in which differences might eventually fade away and be forgotten. The false pretense that we have forgotten them, or can pretend we are in the course of forgetting them, is merely an all too common ideology that sanctions oppression.

Fundamental to the religious reforms for which Kant is hoping, however, would be something Mendelssohn would insist upon at least as much as would Kant – namely, respecting all human beings as regards their freedom of conscience. A spirit of liberality rather than coercive tribalism must come to characterize every faith, as it would do equally for Kant's Christianity and Mendelssohn's Judaism. Only this can pave the way for a universal human community, whether it takes the form of a single united moral community (as Kant hopes) or a plurality of distinct communities (as Mendelssohn would favor). If every community of faith respects the freedom of conscience of its own members, then any remaining differences between Kant's hopes and Mendelssohn's will diminish and will even cease to matter.

Kant's paradigm illustration of historical reform regarding traditional Christian doctrine is his treatment of human sin and salvation through divine grace. This doctrine can be understood in such a way that it leads back to a rational religion acceptable to all on universally valid grounds. Some parts of the traditional Christian doctrine are acceptable religious options, but not essential to a "saving faith," and therefore no barrier to the acceptance in the universal moral community of either Christians or non-Christians who cannot in good conscience accept them. The traditional morally indifferent practices sanctioned by ecclesiastical faith will, Kant hopes, increasingly come to be seen in a similar way. They should not be viewed as genuine service of God that could solicit special divine favor and substitute for conduct that universal reason will approve as good.

Mendelssohn's friend Lessing argued that religion belongs to the education of the human species. Kant sees religion in much the same way. As human beings become enlightened, thinking for themselves by taking responsibility for their own thoughts, feelings, and actions, they will no longer need to see the letter of

traditional scriptures, creeds, catechisms, obedience to ecclesiastical authorities, as part of the true service of God. Kant hopes these reforms will bring institutionalized religion into increasing harmony with science, historical truth, and rational morality, while deepening people's moral lives through the symbolism that has always belonged essentially to all religious thinking. In this way, the boundary between the inner and outer circles of religion will blur: revealed religion will recognize itself as the historical vehicle for pure rational religion, and rational religion will celebrate its greatest successes in the modernizing, enlightened interpretation (*never merely the rejection*) of ecclesiastical tradition and its symbols.

Kant's vision for religious reform may now be hard for us to share. It may even be hard to catch a glimpse of it because it is so obvious that things have not gone the way Kant hoped. Instead of becoming a vehicle of enlightenment, religious traditions have even tended militantly to oppose the very changes Kant is advocating. Popular ecclesiastical faiths have tended to retreat, each into its own tradition. Religion then becomes a vehicle not of a universal human community but of ethnic separateness and bigotry. To some, "religious freedom" comes to mean the right to be intolerant toward outsiders, to insist dogmatically on the most backward and inhuman of one's parochial traditions. Or else "religion" comes to mean the defiance of reason, science, human progress, free and critical communication between people, even the rejection of basic respect for human dignity. Different religious sects then come to agree only in rejecting reason. Each sets up its own border wall of security against a modern world and a secular culture that they regard as hostile. Modernity, enlightenment, even science becomes the enemy. The religious attitude then retreats into delusion, which now becomes a matter of principle with it.

In the famous coinage of Max Weber, the modern world has become *entzaubert*: "disenchanted." Enlightenment reason strikes many people as something requiring of us that we live without religious sources of meaning. Secularism is resigned to living in a disenchanted world; it attempts to find some meaning in a secular culture or politics in which religion no longer has any part to play. On this view, religion must be excluded because it is seen, with justification, as only a source of division, unreason, and intolerance. Hume's essay on *The Natural History of Religion* diagnoses obscurity, absurdity, superstition, bigotry, and intolerance as essential to the appeal of religion. Popular religion at all times and places, perhaps most conspicuously in our age, seems to confirm Hume's harsh judgment.

Kant is aware of the consequences of this side of religion's appeal: delusion, fetishism, counterfeit service of God, priestcraft, violation of freedom of conscience. We have seen that Kant, along with the Bible, is aware of the permanent threat of idolatry: worshipping the symbol in its literal meaning instead of what it symbolizes. This allows the symbol to express uncritically, even triumphantly, reprehensible thoughts and emotions, and it relieves the worshipper of the responsibility for interpreting religious symbols rationally. Despite this, however, Kant views the present historical predicament of religion

with hope: “Should one ask, Which period of the entire church history in our cognizance up to now is the best? I reply without hesitation: *The present*” (R 6:131).

Living in the age when critical biblical scholarship first began to offer an enlightened and critical stance toward traditional scriptures, Kant applauded the way this tendency was opening up religious thought and challenging the use of scriptural authority only to support backward and superstitious attitudes:

In matters which ought to be moral and soul-improving by nature, reason has wrested itself free from the burden of a faith constantly exposed to the arbitrariness of its interpreters . . . It has accepted, *in the first place*, the principle of reasonable moderation in claims concerning anything that goes by the name of revelation . . . [T]hat scripture which, in its practical content contains much that is godly may be regarded (with respect to what is historical in it) as divine revelation (R 6:131–132).

Since we have no *cognitive* awareness of God, or of his actions or teachings, the “godliness” of scriptural narratives and doctrines must be judged by their *moral* content. Thus we can learn from it, and regard it as authoritative for us, only by means of an interpretation of its content that harmonizes with our best and most enlightened understanding of morality. “The *second principle* [therefore] is that since the sacred narrative is adopted only for the sake of ecclesiastical faith, it should be expounded in the interest of morality . . . which is alone capable of making us well-pleasing to God” (R 6:132–133). Kant was hopeful that the enlightened side of religious culture would win out over the side of religion that opposes reason and morality.

Like other enlightenment thinkers and heirs of enlightenment, Kant regards the natural world’s *Entzauberung* as a liberation – a passage through adolescence of the human spirit to its self-confident maturity. Kant also had cautious hopes for politics, the state, and commerce, and also for science, learning, and scholarly communication. But it was in religion, the moral community or church, that he most saw a hopeful future for true human community. “The leading-string of holy tradition, with its appendages, its statutes and observances, which in its time did good service, become bit by bit dispensable, yea, finally, when the human being enters upon his adolescence, turn into a fetter” (R 6:121).

Kant believed that we can retain the emotional power of religion while at the same time taking responsibility for the moral interpretation and rational understanding of its symbols, no longer regarding them for their own sake as the true reality or authority but seeing human reason itself and its morality as the reality symbolized. True religion will then reveal itself as the church universal – the one, true, rational human moral community.

This seems not to be the way things have gone, however. And whatever else we moderns may agree or disagree about, modernity’s most defining feature is self-hatred: *hatred of modernity itself*. We differ among ourselves mainly in our diagnosis of the sickness from which modernity suffers and in our hopes for

a cure. Religious institutions and traditions retain their distinctiveness – not only from each other but from what we now call “secular” culture, which they often see as the enemy, or at best with which they maintain an uneasy truce. Religion is more the guardian of old ways, which it seeks to protect from modernity as from moral corruption. Many secularists, among whom I would number myself, express their discontent by wanting to transform the social world, the political and economic sphere, even to transform it radically – to remake it into a different real (social, economic, political) world. For both groups, it may be impossible to share Kant’s hopes, perhaps difficult even to imagine them. It is easy to see Kant’s hopes for religion only as the hope to reduce religion to, or replace it with, secular morality. This, however, gets Kant exactly wrong. For him religion was the very heart of an enlightenment transformation of human community. Religion was not a way of protecting or reanimating an old world; even more than politics, science, commerce, shared labor, or anything else, religion was to be the way forward to a new and more enlightened world.

Kant’s *Religion* therefore comes at us from an intellectual and spiritual direction we have a hard time understanding. It’s no wonder that Kant’s vision is so often and so badly misunderstood and so widely treated with rejection or condescension. It occupies a place, a philosophical position, a position in modern culture, and it embodies a set of religious hopes, that we now have difficulty even imagining.

I hope this book can help us to imagine them. Those who do may come to agree with me in thinking that if modern culture had gone on to develop in the ways Kant hoped, then we would all be far better off than we are – intellectually, morally, and spiritually. This leads me also to think that if we are ever to find some way forward from where we are, it may benefit us to understand Kant’s religious vision better than we do.

Bibliography

- Adams, Robert M. (1987). *The Virtue of Faith and Other Essays in Philosophical Theology*. New York: Oxford University Press.
- Adams, Robert M. (1999). *Finite and Infinite Goods*. Oxford: Oxford University Press.
- Adams, Robert M. (2018). Introduction to *Religion within the Boundaries of Mere Reason and Other Writings* by Immanuel Kant. Edited and translated by Allen Wood and George di Giovanni. 2nd ed. Cambridge: Cambridge University Press.
- Adickes, Erich (1920). Kants Opus Postumum, Kant-Studien Ergänzungsheft 50.
- Allison, Henry E. (1990). *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Allison, Henry E. (2011). *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press.
- Allison, Henry E. (2018 [1966]). *Lessing and the Enlightenment: His Philosophy of Religion and Its Relation to Eighteenth-Century Thought*. Albany: SUNY Press.
- Alston, William (1996). "Belief, Acceptance, and Religious Faith." In J. Jordan (ed.), *Freedom and Rationality*, pp. 3–27. Lanham, MD: Rowman and Littlefield.
- Altman, Matthew (ed.) (2017). *The Palgrave Kant Handbook*. London: Palgrave Macmillan.
- Anderson-Gold, Sharon, and Pablo Muchnik (eds.) (2010). *Kant's Anatomy of Evil*. Cambridge: Cambridge University Press.
- Anscombe, G. E. M. (1963). *Intention*. 2nd ed. Oxford: Blackwell.
- Aristotle (1999). *Nicomachean Ethics*. 2nd ed. Translated by Terence Irwin. Indianapolis, IN: Hackett.
- Arkush, Allan (1994). *Moses Mendelssohn and the Enlightenment*. Albany: SUNY Press.
- Arendt, Hannah (1994). *The Origins of Totalitarianism*. San Diego: Harcourt Brace.
- Arendt, Hannah, and Karl Jaspers (1992). *Correspondence, 1926–1969*. Edited by Lotte Kohler and Hans Saner. Translated by Robert and Rita Kimber. San Diego: Harcourt Brace.
- Beck, Lewis White (1960). *Kant's Critique of Practical Reason: Commentary*. Indianapolis, IN: Bobbs-Merrill.
- Beiser, Frederick C. (1987). *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, MA: Harvard University Press.

- Beiser, Frederick C. (1996). *The Sovereignty of Reason: The Defense of Rationality in the Early English Enlightenment*. Princeton, NJ: Princeton University Press.
- Berlin, Isaiah (1977). "David Hume and the Sources of German Anti-rationalism." In G. P. Morice (ed.), *David Hume Bicentenary Papers*, pp. 204–235. Austin: University of Texas Press.
- Berlin, Isaiah (2002). *Liberty*. Edited by Henry Hardy. Oxford: Clarendon.
- Bernstein, Leonard (1971). *Mass*. London: Boosey and Hawkes.
- Bernstein, Richard (2002). *Radical Evil: A Philosophical Interrogation*. Cambridge, MA: Polity Press.
- Bielefeldt, Heiner (2001). *Kants Symbolik: Ein Schlüssel zur kritischen Freiheitsphilosophie*. Munich: Alber. [English translation: *Symbolic Representation in Kant's Practical Philosophy*. Translated by Heiner Bielefeldt. New York: Cambridge University Press, 2003.]
- Blackburn, Simon (1998). *Ruling Passions*. Cambridge: Cambridge University Press.
- Byrne, Peter (1979). "Kant's Moral Proof of the Existence of God." *Scottish Journal of Theology* 32, pp. 333–335.
- Card, Claudia (2002). *The Atrocity Paradigm: A Theory of Evil*. New York: Oxford University Press.
- Caswell, Matthew (2007). "Kant on the Diabolical Will: A Neglected Alternative?" *Kantian Review* 12, no. 2, pp. 147–157.
- Chignell, Andrew (2006). "Beauty as a Symbol of Natural Systematicity." *British Journal of Aesthetics* 46, no. 4, pp. 406–415.
- Chignell, Andrew (2007). "Belief in Kant." *Philosophical Review* 116, no. 3, pp. 323–360.
- Chignell, Andrew (2010a). "The Devil, the Virgin, and the Envoy: Symbols of Moral Struggle in Religion, Part Two, Section Two." In Otfried Höffe (ed.), *Kant: Religion innerhalb der Grenzen der bloßen Vernunft*. Klassiker auslegen. Berlin: Akademie Verlag.
- Chignell, Andrew (2010b). "Real Repugnance and Belief about Things-in-Themselves: A Problem and Kant's Three Solutions." In James Krueger and Benjamin Bruxvoort Lipscomb (eds.), *Kant's Moral Metaphysics*, pp. 177–209. Berlin: DeGruyter.
- Chignell, Andrew (2016). "Rational Hope, Possibility, and Divine Action." In Michalson (ed.), pp. 98–117.
- Chignell, Andrew (2018). "Religious Dietary Practices and Secular Food Ethics: Or, How to Hope Your Food Choices Make a Difference Even When You Reasonably Believe They Don't." In Anne Barnhill, Mark Budolfson, and Taylor Doggett (eds.), *Oxford Handbook of Food Ethics*, pp. 287–312. Oxford: Oxford University Press.
- Chignell, Andrew (ed.) (2019). *Evil: A History*. Oxford: Oxford University Press.
- Clifford, William Kingdon (1999). *The Ethics of Belief and Other Essays*. Amherst, NY: Prometheus.
- Conze, Edward (1963). "Spurious Parallels to Buddhist Philosophy." *Philosophy East and West* 13, no. 2 (July), pp. 105–115.
- Coxe, William (1816). *Memoirs of Robert Walpole*. 4 vols. London: Longman, Hurst, Rees, Orme, and Brown.
- Cummiskey, David (2017). "Buddhist Modernism and Kant on Enlightenment." In Steven Emmanuel (ed.), *Buddhist Philosophy: A Comparative Approach*, pp. 205–220. Oxford: Wiley-Blackwell.

- Dalberg-Acton, John (1907). "Letter to Bishop Mandell Creighton," April 5, 1887. In J. N. Figgis and R. V. Laurence (eds.), *Historical Essays and Studies*, pp. 503–508. London: Macmillan.
- Davidson, Donald (1963). "Actions, Reasons, and Causes." *Journal of Philosophy* 60, pp. 685–700.
- Di Cenzo, James J. (2012). *Kant's Religion within the Boundaries of Mere Reason: A Commentary*. Cambridge: Cambridge University Press.
- Dilthey, Wilhelm (1890). "Der Streit Kants mit der Zensur über das Recht freier Religionsforschung." *Archiv für Geschichte der Philosophie* 3, pp. 418–50.
- Doris, John M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Doris, John M. (ed.) (2010). *The Moral Psychology Handbook*. Oxford: Oxford University Press.
- Doris, John M. (2017). The John L. Doris Memorial Lecture: "Marking Good: Can We Realize Our Moral Aspirations?" www.bctr.cornell.edu/tag/john-doris-memorial-lecture/.
- Eagleton, Terry (2011). *Why Marx Was Right*. New Haven, CT: Yale University Press.
- Eliot, Thomas Stearns (1963 [1935]). *Murder in the Cathedral*. New York: Harcourt Brace.
- Fichte, Johann Gottlieb (1970 [1790]). *Aphorismen über Religion und Deismus*. Fichtes *Sammtliche Werke*. Vol. 5. Edited by I. H. Fichte. Berlin: W. deGruyter.
- Fichte, Johann Gottlieb (1970 [1806]). *Anweisung zum seligen Leben, oder auch die Religionslehre*. Fichtes *Sammtliche Werke*. Vol. 5. Edited by I. H. Fichte. Berlin: W. de Gruyter.
- Fichte, Johann Gottlieb (2010 [1793]). *Attempt at a Critique of All Revelation*. Vol. 5 of Fichtes *Sammtliche Werke*. Translated by G. Green. Edited by Allen W. Wood. Cambridge: Cambridge University Press.
- Firestone, Chris L. (2009). *Kant and Theology at the Boundaries of Reason*. Farnham: Ashgate.
- Firestone, Chris L., and Nathan Jacobs (2008). *In Defense of Kant's Religion*. Bloomington: Indiana University Press.
- Firestone, Chris L., and Stephen Palmquist (eds.) (2006). *Kant and the New Philosophy of Religion*. Bloomington: Indiana University Press, 2006.
- Flikschuh, Katrin (2010). "Gottesdienst und Afterdienst: Die Kirche als öffentliche Institution?" In Otfried Höffe (ed.), *Immanuel Kant: Die Religion innerhalb der Grenzen der bloßen Vernunft*, pp. 193–210. Berlin: Akademie Verlag.
- Freud, Sigmund (1964). *The Future of an Illusion*. Translated by W. D. Robson-Scott. Revised by James Strachey. New York: Doubleday Anchor.
- Frierson, Patrick (2003). *Freedom and Anthropology in Kant's Moral Philosophy*. Cambridge: Cambridge University Press.
- Frierson, Patrick (2008). "Empirical Psychology, Common Sense, and Kant's Empirical Markers for Moral Responsibility." *Studies in the History and Philosophy of Science* 34, pp. 473–482.
- Goethe, J. Wolfgang von (1988). *Briefe*. Edited by K. R. Mandelkow. Munich: Deutsche Taschenbuch.
- Greene, Theodore M. (1960). "The Historical Context and Religious Significance of Kant's Religion." Introduction to *Religion within the Limits of Reason Alone* by Immanuel Kant, pp. vii–lxxvii. Translated by Theodore M. Greene and Hoyt Hudson. New York: Harper.

- Grenberg, Jeanine (2005). *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue*. Cambridge: Cambridge University Press.
- Griswold, Charles L. (2007). *Forgiveness: A Philosophical Exploration*. Cambridge: Cambridge University Press.
- Guyer, Paul (2005). *Kant's System of Nature and Freedom*. Oxford: Oxford University Press.
- Guyer, Paul (2015). "Moral Worth and Moral Motivation: Kant's Real View." *International Yearbook for German Idealism* 13, pp. 19–38.
- Hare, John E. (1996). *The Moral Gap: Kantian Ethics, Human Limits, and God's Assistance*. Oxford: Oxford University Press.
- Harman, Gilbert (1999). "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99, pp. 315–331.
- Harman, Gilbert (2009). "Skepticism about Character Traits." *The Journal of Ethics* 13, pp. 235–242.
- Harvey, Van A. (1996). *The Historian and the Believer: The Morality of Historical Knowledge and Christian Belief*. 2nd ed. Chicago: University of Chicago Press.
- Hegel, G. W. F. (1981). *Phenomenology of Spirit*. Translated by A. V. Miller. Oxford: Oxford University Press.
- Hegel, G. W. F. (1991). *Elements of the Philosophy of Right*. Translated by H. B. Nisbet. Edited by Allen Wood. Cambridge: Cambridge University Press.
- Herman, Barbara (2002). "Could It Be Worth Thinking about Kant on Sex and Marriage?" In Louise Antony and Charlotte Witt (eds.), *A Mind of One's Own*. 2nd ed., pp. 53–72. Boulder, CO: Westview.
- Herman, Barbara (2018). "Religion and the Highest Good: Speaking to the Heart of Even the Best of Us." In Kate Moran (ed.) *Kant on Freedom and Spontaneity*, pp. 214–230. Cambridge: Cambridge University Press.
- Hudson, Hud (1994). *Kant's Compatibilism*. Ithaca, NY: Cornell University Press.
- Hume, David (1980). *Dialogues Concerning Natural Religion*. 2nd ed. Edited by Richard Popkin. Indianapolis, IN: Hackett.
- Hume, David (1983). *History of England*. Indianapolis, IN: Liberty Fund.
- Hume, David (1992). *The Natural History of Religion*. Edited with an introduction by James Fieser. New York: Macmillan.
- Hume, David (2007). *Enquiry Concerning Human Understanding*. Edited by David Millican. Oxford: Oxford University Press.
- Insole, Christopher J. (2016). *The Intolerable God*. Grand Rapids, MI: Eerdmans.
- Kahn, Samuel (2018). "Kant's Post-1800 Disavowal of the Highest Good Argument for the Existence of God." *Kant Yearbook* 10, no. 1, pp. 63–83.
- Kahn, Samuel (2019). *Kant, Ought Implies Can, the Principle of Alternate Possibilities and Happiness*. Lanham, MD: Rowman and Littlefield.
- Kail, P. J. E. (2007). "Understanding Hume's Natural History of Religion." *Philosophical Quarterly* 57, no. 227 (April 1), pp. 190–211.
- Kang, Y. A. (1985). *Schema and Symbol: A Study of Kant's Doctrine of Schematism*. Amsterdam: Free University Press.
- Kierkegaard, Søren (1980 [1848]). *The Sickness unto Death*. Translated by Howard Hong. Princeton, NJ: Princeton University Press.
- Kierkegaard, Søren (2006 [1843]). *Fear and Trembling*. Translated by Sylvia Walsh. Cambridge: Cambridge University Press.

- Korsgaard, Christine (1996). *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Lamott, Anne (2005). *Plan B: Further Thoughts on Faith*. New York: Penguin.
- Lessing, G. E. (2005). *Theological Writings*. Edited and translated by H. B. Nisbet. Cambridge: Cambridge University Press.
- Longuenesse, Béatrice (2017). *I, Me, Mine: Back to Kant and Back Again*. Oxford: Oxford University Press.
- Mariña, Jacqueline (1997). "Kant on Grace: A Reply to His Critics." *Religious Studies* 33, pp. 379–400.
- Mariña, Jacqueline (2009). "Making Sense of Kant's Highest Good." *Kant-Studien* 91, no. 3, pp. 329–355.
- Mariña, Jacqueline (2017). "Kant's Robust Theory of Grace." *International Journal of Philosophy* 6, pp. 302–320.
- Matušík, Martin Beck (2008). *Radical Evil and the Scarcity of Hope*. Bloomington: Indiana University Press.
- Mendelssohn, Moses (1969). *Jerusalem and Other Jewish Writings*. Translated and edited by Alfred Jospe. New York: Schocken.
- Mendelssohn, Moses (1983 [1783]). *Jerusalem; Or, On Religious Power and Judaism*. Translated by Allan Arkush. London: Brandeis University Press.
- Mendelssohn, Moses (2007 [1767]). *Phädon; or, On the Immortality of the Soul*. Translated by Patricia Noble. New York: Peter Lang.
- Mendelssohn, Moses (2011 [1785]). *Morning Hours: Lectures on God's Existence*. Translated by Daniel Dahlstrom and Corey Dyck. Dordrecht: Springer.
- Merritt, Melissa (2018). *Kant on Reflection and Virtue*. Cambridge: Cambridge University Press.
- Michalson, Gordon (1986). "The Non-moral Element in Kant's 'Moral Proof' of the Existence of God." *Scottish Journal of Theology* 39, pp. 501–515.
- Michalson, Gordon (1990). *Fallen Freedom: Kant on Radical Evil and Moral Regeneration*. Cambridge: Cambridge University Press.
- Michalson, Gordon (1999). *Kant and the Problem of God*. Oxford: Blackwell.
- Michalson, Gordon (ed.) (2016). *Cambridge Critical Guide to Religion within the Boundaries of Mere Reason*. Cambridge: Cambridge University Press.
- Mill, John Stuart (1978). *On Liberty*. Edited by Elizabeth Rapaport. Indianapolis, IN: Hackett.
- Mill, John Stuart (2001 [1861]). *Utilitarianism*. 2nd ed. Edited by George Sher. Indianapolis, IN: Hackett.
- Montaigne, Michel Eyquem (1991). *The Complete Essays*. Translated by M. A. Screech. London: Penguin.
- Moors, Martin (2017). "The Fate of Religion within the Boundaries of Mere Reason." In Altman (ed.) (2017).
- Morgan, Seiriol (2005). "The Missing Formal Proof of Humanity's Radical Evil in Kant's Religion." *Philosophical Review* 114, no. 1, pp. 63–114.
- Morton, Jennifer M., and Sarah K. Paul (2019). "Grit." *Ethics* 129, no. 2 (January), pp. 175–203.
- Murti, T. R. V. (2008 [1955]). *The Central Philosophy of Buddhism: A Study of the Madhyamika System*. London: Routledge.
- Neuhouser, Frederick (2008). *Rousseau's Theodicy of Self-Love*. Oxford: Oxford University Press.

- Nietzsche, Friedrich (1967). *On the Genealogy of Morals*. Translated by Walter Kaufmann. New York: Vintage.
- Palmquist, Stephen (1992). "Does Kant Reduce Religion to Morality?" *Kant-Studien* 83, no. 2, pp. 129–148.
- Palmquist, Stephen (2000). *Kant's Critical Religion*. Vol. 2 of Kant's System of Perspectives. Aldershot: Ashgate.
- Palmquist, Stephen (2008). "Kant's Quasi-transcendental Argument for a Necessary and Universal Evil Propensity in Human Nature." *Southern Journal of Philosophy* 46 (Summer), pp. 261–297.
- Palmquist, Stephen (2015). *Comprehensive Commentary on Kant's Religion within the Boundaries of Bare Reason*. Oxford: Wiley-Blackwell.
- Papish, Laura (2018). *Kant on Evil, Self-Deception, and Moral Reform*. Oxford: Oxford University Press.
- Pascal, Blaise (1995). *Pensées*. Translated by A. J. Krailsheimer. London: Penguin. Cited by section (§).
- Pasternack, Lawrence (2014). *Guidebook to Kant's Religion within the Boundaries of Mere Reason*. London: Routledge.
- Pasternack, Lawrence (2017a). "Kant on Faith, Religious Assent, and the Limits of Knowledge." In Altman (ed.).
- Pasternack, Lawrence (2017b). "The 'Two Experiments' in Kant's Religion: Dismantling the Conundrum." *Kantian Review* 22, pp. 107–131.
- Pasternack, Lawrence, and Philip Rossi (2014). "Kant's Philosophy of Religion." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014 edition). <https://plato.stanford.edu/archives/fall2014/entries/kant-religion/>.
- Plato (1980). *Gorgias*. Translated by Terence Irwin. Oxford: Clarendon.
- Prichard, H. A. (2002) *Moral Writings*. Edited by J. MacAdam. Oxford: Clarendon.
- Pufendorf, Samuel (2005 [1729]). *De jure naturae et gentium (On the Law of Nature and of Nations)*, Fourth Edition, . . . To Which Is Now Prefixed Mr. Barbeyrac's Prefatory Discourse, . . . Done into English by Mr. [George] Carew, Buffalo, NY, William S. Hein; reprint of London: S. Aris.
- Quinn, Philip (1978). *Divine Commands and Moral Requirements*. Oxford: Clarendon Press.
- Ray, Matthew Alun (2003). *Subjectivity and Irreligion: Atheism and Agnosticism in Kant, Schopenhauer, and Nietzsche*. Aldershot: Ashgate.
- Raz, Joseph (2010). "On the Guise of the Good." In S. Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. Oxford Scholarship Online.
- Reath, Andrews (1988). "Two Conceptions of the Highest Good in Kant." *Journal of the History of Philosophy* 26, no. 4 (October), pp. 593–619.
- Reinhold, Karl Leonhard (2005). *Letters on the Kantian Philosophy*. Edited by K. Ameriks. Cambridge: Cambridge University Press.
- Rosenzweig, Franz (1985 [1930]). *The Star of Redemption*. Translated by Willam Hallo. South Bend, IN: Notre Dame University Press.
- Ross, W. D. (2002). *The Right and the Good*. Edited by Philip Stratton-Lake. Oxford: Oxford University Press.
- Rossi, Philip, and Michael Wreen (eds.) (1991). *Kant's Philosophy of Religion Re-examined*. Bloomington: Indiana University Press.
- Rousseau, Jean-Jacques (1997). *The Discourses and Other Political Writings*. Edited by Victor Gourevitch. Cambridge: Cambridge University Press.

- Rowe, William L. (1968). *Religious Symbols and God: A Philosophical Study of Tillich's Theology*. Chicago: University of Chicago Press.
- Royce, Josiah (2008 [1908]). *The Philosophy of Loyalty*. London: Macmillan; Reprint: BiblioLife.
- Sala, G. B. (2004). "The Christology in Kant's Religion." *Rivista di Filosofia Neo-Scholastica* 96, no. 2-3, pp. 235-304.
- Sartre, Jean-Paul (1956). *Being and Nothingness*. Translated by Hazel Barnes. New York: Philosophical Library.
- Savage, Denis (1988). "Kant's Rejection of Divine Revelation and his Theory of Radical Evil." In Philip J. Rossi and Michael Wreen (eds.), *Kant's Philosophy of Religion Reconsidered*, pp. 54-75. Bloomington: Indiana University Press.
- Schapiro, Tamar (2006). "Kantian Rigorism and Mitigating Circumstances." *Ethics* 117, no. 1, pp. 32-57.
- Schleiermacher, Friedrich Daniel Ernst (1998). "Review of Immanuel Kant's Anthropology from a Pragmatic Point of View (1799)." From *Athenaeum* 2, no. 2, translated by Jacqueline Mariña, *New Athenaeum*, pp. 11-31.
- Schopenhauer, Arthur (1958). *The World as Will and Representation*. Vol. 1. Translated by E. J. F. Payne. New York: Dover.
- Schott, Robin May (1988). *Cognition and Eros: A Critique of the Kantian Paradigm*. Boston: Beacon.
- Schrift, A. D. (ed.) (2005). *Modernity and the Problem of Evil*. Bloomington: Indiana University Press.
- Schweitzer, Albert (2005 [1911]). *Quest for the Historical Jesus*. Translated by W. Montgomery. Mineola, NY: Dover.
- Setiya, Kieran (2007). *Reasons without Rationalism*. Princeton, NJ: Princeton University Press.
- Shantz, Douglas H. (2013). *An Introduction to German Pietism: Protestant Renewal at the Dawn of Modern Europe*. Baltimore, MD: Johns Hopkins University Press.
- Sidgwick, Henry (1981). *Methods of Ethics*. 7th ed. Indianapolis, IN: Hackett.
- Silber, John (1960). "The Ethical Significance of Kant's Religion." Introduction to *Religion within the Limits of Reason Alone* by Immanuel Kant, pp. tk-tk. Translated by Theodore M. Greene and Hoyt Hudson. New York: Harper.
- Silber, John (1991). "Kant at Auschwitz." In G. Funke and T. M. Seebohm (eds.), *Proceedings of the Sixth International Kant Congress*, pp. 177-211. Washington, D. C.: Center for Advanced Research in Phenomenology and University Press of America.
- Silber, John (2012). *Kant's Ethics*. Boston: De Gruyter.
- Sinclair, Upton (1934). *I, Candidate for Governor, and How I Got Licked*. Berkeley: University of California Press.
- Smith, Adam (2000). *The Theory of Moral Sentiments*. Amherst, NY: Prometheus.
- Spinoza, Baruch (1982 [1677]). *Ethics and Selected Letters*. Translated by Samuel Shirley. Indianapolis, IN: Hackett.
- Stern, Robert (2004). "Does 'Ought' Imply 'Can'? And Did Kant Think It Does?" *Utilitas* 16, no. 1, pp. 42-61.
- Stoeffler, F. Ernest (1973). *German Pietism during the Eighteenth Century*. Studies in the History of Religion 24. Leiden: E. J. Brill.
- Sussman, David (2001). *The Idea of Humanity: Anthropology and Anthropolonomy in Kant's Ethics*. New York: Routledge.

- Sussman, David (2005). "Perversity of the Heart." *Philosophical Review* 114, no. 2, pp. 153–177.
- Sussman, David (2009). "For Badness' Sake." *Journal of Philosophy* 106, no. 11, pp. 613–628.
- Swinburne, Richard (2004). *The Existence of God*. 2nd ed. Oxford: Oxford University Press.
- Tenenbaum, Sergio (2007). *Appearances of the Good: An Essay on Practical Reason*. Cambridge: Cambridge University Press.
- Thorndike, Oliver (ed.) (2011). *Rethinking Kant*. Vol. 3. Cambridge: Cambridge Scholars.
- Tillich, Paul (1951, 1957, 1963). *Systematic Theology*. 3 vols. Chicago: University of Chicago Press.
- Tillich, Paul (1952). *The Courage to Be*. New Haven, CT: Yale University Press.
- Tillich, Paul (1957). *Dynamics of Faith*. New York: Harper and Row.
- Velleman, J. David (1992). "The Guise of the Good." *Noûs* 26, pp. 1–26.
- Ware, Owen (2009). "The Duty of Self-Knowledge." *Philosophy and Phenomenological Research* LXXIX, no. 3, pp. 671–698.
- Watkins, Eric (2005). *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press.
- Williams, Bernard (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
- Willis, Janice Dean (2008). *Dreaming Me: Black, Baptist, and Buddhist*. Sommerville, MA: Wisdom.
- Wolterstorff, Nicholas (1991). "Conundrums in Kant's Rational Religion." In Philip J. Rossi and Michael Wreen (eds.), *Kant's Philosophy of Religion Reconsidered*, pp. 40–53. Bloomington: Indiana University Press.
- Wolterstorff, Nicholas (1998). "Is It Possible and Desirable for Theologians to Recover from Kant?" *Modern Theology* 1, no. 1, pp. 1–18.
- Wood, Allen (1970, 2009). *Kant's Moral Religion*. Ithaca, NY: Cornell University Press.
- Wood, Allen (1978, 2009). *Kant's Rational Theology*. Ithaca, NY: Cornell University Press.
- Wood, Allen (1984). "Kant's Compatibilism." In Allen Wood (ed.), *Self and Nature in Kant's Philosophy*, pp. 73–100. Ithaca, NY: Cornell University Press.
- Wood, Allen (1992). "Rational Theology, Moral Faith, and Religion." In Paul Guyer (ed.), *Cambridge Companion to Kant*, pp. 394–417. New York: Cambridge University Press.
- Wood, Allen (1999). *Kant's Ethical Thought*. New York: Cambridge University Press.
- Wood, Allen (2002). *Unsettling Obligations: Essays on Reason, Reality, and the Ethics of Belief*. Stanford: CSLI.
- Wood, Allen (2005). *Kant*. Oxford: Wiley-Blackwell.
- Wood, Allen (2008). *Kantian Ethics*. New York: Cambridge University Press.
- Wood, Allen (2010). "Ethical Community, Church, and Scripture." In O. Höffe (ed.), *Kant: Religion innerhalb der Grenzen der bloßen Vernunft*. Klassiker auslegen. Berlin: Akademie Verlag.
- Wood, Allen (2012). "Recht und Universität bei Kant." In Ludger Honnefelder et. al. (eds.), *Kants Streit der Fakultäten: Oder der Bildung zwischen Lebenswelt und Wissenschaften*, pp. 54–76. Berlin: Berlin University Press.

- Wood, Allen (2014). *The Free Development of Each: Studies on Freedom, Right, and Ethics in Classical German Philosophy*. Oxford: Oxford University Press.
- Wood, Allen (2015). "Kant on the History of Ethics." In O. Sensen and L. Denis (eds.), *Kant's Lectures on Ethics: A Critical Guide*, pp. 120–137. Cambridge: Cambridge University Press.
- Wood, Allen (2017). *Formulas of the Moral Law*. Kant Elements. Cambridge: Cambridge University Press.
- Wood, Allen (2019). "Evil in Classical German Philosophy: Evil, Selfhood, and Despair." In Chignell (ed.), pp. 322–348.
- Zuckert, Rachel (2010). "Kant's Account of Practical Fanaticism." In Benjamin Lipscomb and James Krueger (eds.), *Kant's Moral Metaphysics: God, Freedom, and Immortality*, pp. 291–317. Berlin: De Gruyter.

Index

An ‘n’ following a page number indicates a footnote.

- abortion, xvi
- Abraham and Isaac, 13–14
- acceptance (*Annehmen*), 54, 55, 57
- acedia, 67
- action and agency. *see also* choice; deeds; ends;
 - free action; imputability; legality,
 - legalism and legal metaphors;
 - “rigorism,” Kant’s
- assent and, 31, 56–57, 210
- possible according to reason, 31, 47, 47n15
- Acton, Lord, 35n9
- Adam, 89n1
- Adams, Marilyn McCord, xix
- Adams, Robert M., xix, 29, 41n11, 56n18,
 - 64n2, 133, 151n7
- addiction (*Sucht*), 79
- Adickes, Erich, 34n8
- admiration, 43n13, 80
- advocate (*Sachverwalter*), 126, 134
- aesthetics, Kantian, 1, 5, 6, 120, 172–174. *see also* beauty; sublimity
- “afraid of God,” 16
- agent-attitudes, 43–44, 44n13
- agents, ourselves as, 156
- agnosticism, theoretical, 59
- akrasia*, 67
- Allison, Henry, 10n10, 40, 41, 43, 44, 64n2,
 - 87n25, 93
- ambition, 78n23, 80
- analogies, 5, 52–53, 117–126, 206. *see also*
 - symbols, religious
- anathema, 14, 14n15, 26, 161, 170
- anger at parent example, 145–146
- animal nature, 67, 68, 69, 70–71, 79, 81, 171
- animals, 118–119, 120, 173
- Anscombe, Elizabeth, 96
- Anselm, 27n1
- anthropology, 9, 86–87, 102
- Anthropology* (Kant), 65n3, 87
- anthropomorphism
 - defined, 6
 - literalism and, 7–8, 120
 - polytheism and, 3n3
 - rational and revealed religion compared, 14,
 - 162, 173–174
 - symbolism and, 120, 122, 123, 127
 - “well-pleasing to God” and, 127
- antievindentialism, 56–61
- antinomies, 3n3, 180n5
- antinomy of faith and works, 178–182
- antireligious attitudes, xix
- anti-Semitism, 7n8, 192
- apartheid, 142
- appraiser-attitudes, 43–44, 44n13
- approval of others, 80, 207
- a priori concepts, ideas, and proofs, 5, 52–53,
 - 84–87, 90–91, 117–118, 119–120
- aptitude (*Fertigkeit*), 58
- Aquinas, Thomas, 27n1, 29
- arbitrium sensitivum sed liberum* (sensitive but
 - free power of choice), 65
- Arendt, Hannah, 70
- arguments *ad absurdum practicum*, 42–43, 44
- Aristotle, 31, 49, 72
- Arkush, Allan, 201
- arrogance, 72
- assent (*Fürwahrhalten*)
 - Belief/belief and, 51, 52–60

- assent (*Fürwahrhalten*) (cont.)
 defined, 56–57
 evidentialism and, xx
 to God's existence, 30, 54
 moral progress and, 87
 nonrational, 41n11
 overview, 52–53
 possibility and, 57
 practical grounds of, 29–31
 tacit, 57
 theoretical, 30, 31–32, 34, 46, 47, 51, 53
 third Critique and, 50–56
 three Critiques and, 51
 “assume,” (*annehmen*), 54
 atheism, 7–8, 28n4, 199
 atonement, 126, 132, 134–135, 136, 137,
 137n13, 142, 166–167, 183, 187
 atonement, vicarious, 137n13, 138n16,
 143–148, 144n3
Attempt at a Critique of All Revelation (Fichte),
 20, 21
 attitude of mind (*Denkungsart*), 106–107
 Augustine and Augustinianism, 27n1, 62,
 71n12, 140, 151, 155, 161n12,
 161–163, 178
 authority, 25–26, 121n3, 121–122, 190, 196,
 213. *see also* coercion; conscience;
 freedom of; scripture; service of God;
 states, political; tradition
 autonomy, 40, 44, 172
 avarice, 77, 78n23, 79
 awe, 3

 bad faith, 113
 bad heart, 132–133, 143, 145, 145–146
 badness, 7n8
 banality (of evil), 70
 baptism, 185
 Barth, Karl, xvii
 beauty, 1, 119, 122. *see also* aesthetics, Kantian
 Beck, Lewis White, 1n1, 48, 49
 belief (*Glaube*) (belief or faith). *see also* assent;
 church and ecclesiastical faith;
 conviction; despair; evidentialism; self-
 deception
 advantage or benefit from, 30, 30n6, 41,
 57–58, 207
 Belief compared, 33, 53–54, 56, 59
 evidentialism and xx
 God's commands and, 14–15
 hope and fear *versus*, 197–198
 moral incentives and, 42
 reason and, 201–202, 210
 religious people and, 3, 54–55
 rewards and, 207, 208
 as subjective, 53–56
 Belief, moral (*Glaube*). *see also* assent; moral
 faith in God; symbols
 belief compared, 33, 53–54, 56, 59
 as disposition, 57
 doctrinal, 57
 doubting faith and, 55
 evidentialism and, 56n18
 grace and, 152
 hope and, 51–52, 52n17
 moral argument and, 34, 47
 moral incentives and, 42
 possibility of highest good and, 34–35
 practical, subjective grounds and, 56
 pragmatic Belief compared, 34
 rational and revealed religion compared,
 111, 152
 rational religion and, 128, 145, 152
 subjective grounds and, 52–54
 symbols and, 125
 “well-pleasing to God” and, 127
 ben-David, Lazar, 192n4
 beneficence, 143, 193
 benefit or advantage from believing, 30, 30n6,
 41, 57–58
 benevolence, 198
 Berlin, Isaiah, 59
Berlin Monthly (*Berlinische Monatschrift*,
 Biester, ed.), 20–21
 Bernhard, Isaak, 191
 Bernstein, Leonard, 109n12, 186n2
 Bernstein, Richard, 62n1, 67n5, 85n24, 87n26
 betrayal of trust, 70n11
 Bible, 24, 71n12, 74, 89n1, 89–90, 117, 150,
 176–177, 213. *see also* Paul, Saint;
 Psalm 59; scripture
 biblical scholarship, 214
 Bielefeldt, Heiner, 5n6
 Biester, J. E., 20–21
 bigotry, 213
 bitterness, 137n14
 Blackburn, Simon, 65n3
 blame. *see also* appraiser-attitudes; guilt
 (*Verschuldung*) (third difficulty);
 imputability
 conscientious belief and, 206
 duty to act from duty and, 106
 evil and, 66
 forgiveness and, 137n14

- predispositions and, 70
 punishment and, 135–136
 rational moral incentive and, 43
 recognition of miracles and, 121n3
 time and, 85
 Blanc, Louis, 110n15
 blasphemy, 162n13
 “born again” experiences, 9, 110
Bösewicht (scoundrel), 42, 43
 brain science, 113
 bribe that 7 is not prime, example, 41n11
The Brothers Karamzov (Dostoevsky), xvi
 brutishness, 71
 Buddhism, 3, 3n3
 Buddhists, xix
 Byrne, Peter, 30

 Calvinism, 162n13
 Card, Claudia, 62n1, 68n7, 70n11, 73n14
 caring, 119, 143
 Caswell, Matthew, 69n8
 catechisms, 186, 190, 213
 categorical imperative, 39–40
 Catholicism, xix, 185n1, 188n3, 212
 causality
 evil and, 67
 free action and, 63–64, 95–96, 155–163
 free will and, 65, 65n3, 94–95
 “from itself,” 65
 God and, 66, 155–163, 156n10, 158n11, 181
 guilt and, 144n3
 inclinations and, 93
 natural, 156–157
 objects of loyalties and, 36
 possibility of highest good and, 39
 reasons and, xx, 96n4
 censorship, 11, 20–26
 certainty, 55, 56, 59, 60, 111, 113, 204, 205, 205n18
 change of heart (*Herzensänderung*). *see also*
 ethical community; goodness of heart;
 “old man” and “new man”; pure
 disposition; struggle against evil; time
 acting from duty and, 94, 98–101, 104–110
 atonement and, 126, 137
 characterized, 105
 evil and, 88, 89–90, 104–105
 God and, 151n7, 152, 157
 goodness of heart and, 91–94
 grace and, 151, 157–158, 159–161
 “*Herzensänderung*” and, 110
 hope for, 128, 130
 “justification” and, 138
 legality and, 101–104, 109
 moral incentive and, 89–98, 106, 108–109, 110–111
 “other” *versus* “better” human being and, 153–154
 “Ought implies Can” and, 161
 overviews, 89–90, 117
 prayer and, 186
 rational moral incentive and, 45
 rational religion and revealed doctrine
 compared, 159
 receptivity and, 144, 180
 self-knowledge and, 109n12
 as service of God, 187
 Son of God and, 109–110, 115, 126–127, 130, 138n16
 as symbol, 125–126, 148
 symbolized, 152–153
 time and, 89–90, 93–94, 105–106, 110–114, 128n8, 128–129, 130, 131, 146–147
 vicarious atonement and, 146
 character, human, 73, 96, 98, 100n9, 101–102, 104, 107, 108, 112. *see also* “kind of person”
 chastisement, 50
cherem, 197
 Chignell, Andrew
 Belief and, 33, 51
 on grace and maxims, 162n13
 hope and, 51n16
 on hope and faith, 138n15
 on responsibility for good will, 159
 Stoic maxim and, 154n9
 on struggle against evil, 117n2
 on symbolism, 4n4, 5n6
 “two perspective” problem and, 128n8
 weak moral motivation and, 41n11
 “childish things,” 182–183
 choice, 32–33, 97. *see* free action; incentives
 choice (*Willkür*), 30, 95, 107
 Christ. *see* Jesus Christ
 Christianity. *see also* church and ecclesiastical
 faith; grace and salvation *and other*
 doctrines; Kierkegaard, Søren;
 Lutheranism *and other branches*;
 monotheism; scripture; Wöllner, Johann
 Christoph and his charges
 bad heart and, 146
 change of heart and, 106, 110
 concentric circles image and, 24
 emotions and, 6

- Christianity (cont.)
 experimental testing and, 139–140
 faith and works antinomy and, 180–181
 historical dynamic and, 171
 imposed on others, 212
 intellectual honesty and, 176
 Judaism compared, 191, 203, 203n15
 “justification “ and, 138
 Kant and, xiv–xv, xxi, 1, 157n10, 170,
 181–182, 202
 liberalism and, 21
 mercy and, 142
 modernity and, 20
 morality and, 12, 14, 14n15
 Prussian law and, 192n4
 rational and revealed religion compared,
 25n20
 rational morality and, 14, 14n15
 rational religion and, 181–182
 recommended by Kant, 146
 reform and, 203, 212
 reform of, 181–184, 203, 211, 212
 symbols and, 149–151, 150–151
 toleration and, 11
 white evangelical, xvii–xix
- Christ image, 72n13
- Christology, 116, 116n1
- church and ecclesiastical faith. *see also*
 authority; Christianity *and other faiths*;
 coercion; conscience, freedom of;
 reform; revelation; scripture; service of
 God; tradition
- contemporary, 187, 213
- diversity of, 2, 19, 212
- doubt and, 56n18
- ethical community and, 18, 122, 168–172,
 187, 189, 211–212
- Kant’s hope and, 214
- Mendelssohn on, 192–197, 194n7
- modernity and, 9–10
- moral improvement and, 10, 18
- morality and, 202, 214
- need for, 170–172, 174, 187
- rational and revealed religion compared, 17,
 19, 122–123, 178, 211–212
- reason and, 25–26, 181
- religion compared, 2
- rights and, 197
- state and, 194n7, 194–195, 198–199
- tax-support and, 194n7, 196
- third Critique and, 58
- unity/universality and, 168, 175, 192, 214
- universities and, 25
- visible *versus* invisible, 168
- civilization, 81, 86
- civil rights, 196–197
- clergy-laity distinction, 182–183
- Clifford, W. K., xx, 56, 57–59, 60, 198, 201
- climate disruption (global warming), 36, 52
- coercion, 14, 21, 23, 165–166, 168–169, 186,
 193n6, 193–203, 196n10, 196–199,
 206–207, 212
- cognition (*Erkenntnis*). *see also* empirical
 evidence and experience
- convictions and, 198
- creeds and, 186
- divine *versus* natural causality and, 157n10
- evil and, 66–67
- faith and works and, 179
- freedom and, 160
- of God, 15, 51, 119, 121, 124, 173, 175, 214
- God’s, 94, 174
- grace and, 134, 157n10
- inner feeling and, 177
- Judaism and, 200
- miracles and, 121n3
- overview, 117–118
- possibility of highest good and, 34, 48–49
- practical
- rational habituation and, 96n5
- rational religion and revelation compared, 5
- revelation and, 13, 13n13, 202
- the supernatural and, 156
- the supersensible and, 168
- symbols and, 5, 118, 120–121,
 173, 174
- territories and, 13
- theoretical, 5, 13n13, 15, 28, 33–34
- things-in-themselves and, 65
- coldness of heart, 91–92
- Collegium Fredericianum, 9
- comforting certainty, 59
- comforting complacency, 59
- commands. *see* God’s commands
- commands, unjust, 21, 23
- commands of common legislator, 167–168
- commitments, 210
- common good, 166
- communication, 213. *see also* censorship;
 conscience, freedom of; freedom of
 speech
- communion, 185
- community of rational agents, 82–83, 105. *see also* ethical community

- comparison of self with other people, 71, 76, 78, 78n23, 80
- compatibilism, 65n3
- competitiveness, 77, 78, 79, 80
- complacency, 59, 131
- concentric circles image, 11–13, 19–20, 24, 163, 211. *see also* experimental testing; rational religion and revealed doctrine compared
- concepts, 13
- a priori, 5, 90–91, 117–118
- concepts without objects, 65
- concupiscence (*concupiscentia*, *Concupiscenz*), 71, 71n12, 72
- confidence, 59
- Conflict of the Faculties* (Kant), 22, 24, 25n20, 153
- conscience. *see also* conscience, freedom of certainty and, 205n18
- community and, 208–209
- defined, 205
- duty of self-examination and, 102, 102n10
- ethical community and, 166
- God's judgment and, 206
- individualism and, 208–209
- judgment and, 204n17, 204–205
- Kant and Mendelssohn compared, 203
- Kant on, 204–209
- vicarious atonement and, 143
- violation of, 207
- conscience, freedom of, 22, 190, 192, 197–200, 203, 206–208, 212, 213. *see also* censorship; coercion
- conscious, 190
- conservatism, xvii
- consolation, 55
- constitutions, ethical *versus* political, 169
- constitutions, republican, 183–184
- contemporary culture, 187, 188. *see also* modernity
- contemporary politics, xvii–xix, xxi, 145, 184, 207, 215
- contemporary religion, 213–214
- contempt, 43, 79n23, 92. *see also* appraiser-attitudes
- contracts, 193–194, 195
- conversions, 9, 151. *see also* change of heart conviction (*Überzeugung*), xvi, 41n11, 52–53, 53, 174, 194, 197–199, 201. *see also* evidentialism
- corruption, xx, 41n11, 68. *see also* Henry VIII example
- “cosmic symmetry,” 38
- cosmopolitanism, 87, 211
- Council of Diospolis, 62
- Cranz, August Friedrich, 200n11
- credulity (*Leichtgläubigkeit*), 55, 60
- creeds, 186, 190, 213
- Critique of the Power of Judgment* (Kant), 13, 45, 46n46, 49–51, 58. *see also* aesthetics, Kantian; three Critiques compared
- Critique of Practical Reason* (Kant), 34, 45–49, 58, 76, 91, 103, 117. *see also* three Critiques compared
- Critique of Pure Reason* (Kant), xv, 15, 34n8, 39–45, 47n15, 202. *see also* three Critiques compared
- cruelty, 70n11, 79, 86
- culture (*Kultur*), 49, 77, 79, 80–81. *see also* contemporary culture
- customs (*Sitten*), 172, 202–203
- Cyrus (Persian emperor), xvii–xvii
- Davidson, Donald, 96n4
- death, 131, 149
- debts (*Schulden*), 131–135, 136, 147–149
- decency, 172
- deception, 70n11. *see also* self-deception
- decorum (*Anständigkeit*), 172
- decree of grace, 134–135, 137, 140–141, 144, 146, 147, 152, 162, 164
- deeds, 72, 128, 129. *see also* action and agency
- degradation, 81
- deism, 6, 10n10, 18n17
- deistic concept of God, 119–120, 123–124
- delusion (*Wahn*). *see also* self-deception
- change of heart and, 130–131
- Christianity and, 181
- contemporary, 213
- counterfeit service and, 189
- defined, 7, 80, 121
- examples of, 70n11, 83, 124
- idolatry and, 7, 122, 122n4, 124
- Pietism and, 113
- prayer and, 186
- demonization, 69
- depravity or perversity (*Bösartigkeit*, *Verderbtheit*, *vitiositas*, *pravitas*, *perversitas*, *corruptio*), 83, 84, 88, 89
- Descartes, René, 29, 52, 57, 79
- design, arguments from, 27n1
- design, divine purposive, 10
- desire or sentiment, 41, 71, 72, 92, 119, 172. *see also* happiness; psychology

- despair, 59, 130, 131, 145, 146n4, 209
 devotion, 119
 “diabolical” *versus* “angelic,” 73
 diabolical vices, 69, 69n8, 70n11, 79
 Di Censo, James J., 20n18
 Di Giovanni, George, 20n18
 dignity (rational nature as end-in-itself), 15, 79,
 106, 147n5, 166n2, 213
 Dilthey, Wilhelm, 20n18
 disapproval, 43. *see also* appraiser-attitudes
 discipline, culture of, 49–50. *see also* moral
 discipline
Discourse on the Origin of Inequality among
Men (Rousseau), 76
 disenchantment, 213
 disobedience, 61–62
 disposition (*Gesinnung*). *see also* goodness of
 heart; moral improvement; “old man”
 and “new man”; “prototype of moral
 disposition in its entire purity”
 (formulation of ideal of humanity);
 “well-pleasing to God”
 acting from duty or character and, 98, 104
 assent to belief and, 53, 56–59
 change of heart and, 93, 94, 97, 109, 125
 duty to act from duty and, 107
 empirical experience and, 66
 ethical legislation and, 167
 faith and, 129
 freedom and, 96
 good, 115, 128
 guilt and, 132–133
 improvement and, 93
 as incentive, 97
 morality and, 101–102
 not indifferent, 74
 noumenal world and, 130n10
 pure, 93, 109, 111, 126–131, 138n16,
 143–145, 147–149, 151–152
 pure moral prototype, 116–117
 purity of, 126–127
reality of Son of God and, 138n16
 slavish, 80
 time and, 94, 96, 129
 virtuous, 92
 the divine, 4
 divine justice, 134, 135, 135n11, 143
 divine perfections, 4n4, 15–16, 28, 143
 doctrine, 14, 14n15, 19, 25, 57, 163. *see also*
 grace and salvation *and other doctrines*
 dogmatics, moral, 61, 88
 dogmatism, religious, 25–26
 Dohm, Christian Wilhelm von, 192n5, 196n9,
 196–197
 domain (*Gebeit, ditio*), 13
 domination, 80, 81
 Dostoevsky, Fyodor, xvii, xx
 doubt, 51, 53, 55, 56, 56n18, 59
 doubting faith (*Zweifelsglaube*), 56, 60, 127
Dreaming Me: Black, Baptist, and Buddhist
 (Willis), xix
 duty. *see also* experimental testing; God’s
 commands
 to act from duty (narrow or perfect), 104,
 128n8
 to act from duty (wide or imperfect), 106,
 107–110
 to be beneficent, 99, 100, 100n9
 caring for and, 92
 to choose (narrow), 95n3, 95–96
 coercion and, 196n10
 end of morality and, 33
 God’s commands and, 14–16
 habitual, 92
 happiness and, 39, 99
 highest end and, 39, 48
 of love, 142
 moral and nonmoral incentives and, 108
 moral incentive and, 39
 narrow or perfect, 104, 105, 128n8
 nonmoral motives and, 74
 to ourselves, 119, 121, 143
 perfect and imperfect, 104, 193n6, 193–194
 to preserve one’s life, 99, 100
 to promote one’s own happiness, 99, 100
 rational and revealed, 12–13
 religion and, 2, 3
 religious, 195
 results and, 90
 of self-examination, 102, 103
 self-regarding, 78–79
 surplus good and, 131–132
 symbolism and, 122
 toward beauty or nonrational animals,
 118–119
 toward God, 118–119, 195
 wide, 141–142
 duty, acting from (*aus Pflicht*). *see also*
 incentives, moral and nonmoral
 change of heart and, 94, 98–101, 104–110
 freedom and, 97–98, 109
 in *Groundwork*, 98–102
 legality/morality and, 94
 narrow, 95–96, 96n3

- duty, conformity to (*pflichtmässig*). *see also*
incentives, moral and nonmoral; legality
and legalism
duty to choose (narrow duty) and, 95n3,
95–96
experience and, 90
good disposition and, 129n9
Groundwork and 98–102, 100n9
legality/morality and, 94
supreme good and, 37
- eagle of Jupiter, 6
- Eagleton, Terry, 59–60
- ecclesiastical faith. *see* church
- education, religious, 153, 154–155, 194, 197,
212–213
- election, 162, 162n13
- Eliot, T. S., 113n16
- elites, 79n23
- emotions
analogy and, 5
intellectual conviction and moral resolve
versus., xix
Judaism, Christianity, and Islam and, 6
moral law and, 172
moral motivation and, 174
reason and, 3n3
religion and, 10, 214
revivalism and, 9
symbols and, 5, 7n8, 120, 173n4, 177–178
theistic arguments and, 28
- empirical (psychological) evidence and
experience. *see also* feelings and other
empirical factors; psychology; revelation
- assent and, 52–53
change of heart and, 111–112
dispositions and maxims and, 66
evil maxims and, 74–75
free action and, 63–64, 78
God and, 4, 5, 15, 17, 28, 34
miracles and, 121n3
moral incentive and, 40–42, 41nn11, 12,
43, 95
moral law and, 173
motivation and, 95
origins of morality and, 172
principle of morality and, 90
radical evil and, 87
rational religion and revealed religion
compared, 139
Son of God and, 150
symbols and, 4, 4nn4–5, 123–124, 151
- unsociable sociability and, 85n25
- encouragement, 74
- the end justifies the means*, xviii
- The End of All Things* (Kant), 21
- end of the world, 195
- ends, xviii, 31, 33–34, 40, 49–51, 66, 117. *see also* Formula of the Realm of Ends;
incentives; “whole end that is natural for
every human being”
- enlightened Judaism, 200–201
- enlightenment
contemporary traditions and, 213
defined, 8
ethical community and, 189
learned public and, 23
rational and revealed religion compared, 177,
184, 187
religion and, xxi, 19–20, 186, 215
scripture and tradition and, 177,
212–213, 214
symbols and, 121–122
thinking for oneself and, 25, 121,
123
- Enlightenment (*Haskalah*), 190
- Enlightenment, Scottish, 3n3
- the Enlightenment, 1, 9, 176, 208, 213
- ens rationis*, 65
- ens realissimum*, 4n4, 6, 15–16, 28, 50,
116, 120
- enthusiasm (*Schwärmerei*), 111, 124, 124n5,
156, 168, 172, 190
- “Enthusiasmus,” 124n5
- envy, 69, 70n11, 71, 77, 79
- Epicureans, 38
- Epistle to the Romans* (Barth), xvi
- equality, 182–183, 185
- ethical community. *see also* church and
ecclesiastical faith; individualism;
reform
described, 166–168
emotion and, 3n3
as free, 168–169
highest good and, 38, 167
ideal, 16, 166, 168
juridical *versus*, 165–166
laws and, 167–168, 169
“lost” causes and, 35
morality and, 10
as numerically one, 168, 169–170,
183n6, 203
political community *versus*, 183n6
as pure, 168, 169, 187

- ethical community (cont.)
 rational religion and revealed doctrine
 compared, 122–123, 169–170
 revelation and, 16–17
 scripture and, 174–175
 state of nature and, 164–166
 symbols and, 188
 ethicotheology, 50
 ethnicity, 212, 213
 eudaimonism, 40, 44, 202. *see also* happiness
 eudaimonism and heteronomy, 40
 “Every man has his price,” (Walpole’s thesis),
 75, 102–103, 109
 evidence, moral, 29n4
 evidentialism
 Biester journal’s article and, 21
 characterized, xx
 freedom of conscience and, 197–200
 Mendelssohn and Kant compared,
 201–202, 203
 Mendelssohn on, 198–199, 201–209
 moral faith in God and, 56n18, 56–61
 overview, xx
 evil, propensity to. *see also* self-deception;
 struggle against evil; unsocial sociability
 bad heart and, 143
 change of heart and, 88, 89–90, 104–105, 107,
 109, 117
 commentators and, 84–85
 duty to self-examination and, 102
 freedom and, 95
 guilt and, 148
 innateness of, 62–63, 70, 74, 84–88
 moral *versus* nonmoral incentives and, 83, 93,
 159, 161
 Original Sin and, 62–63
 overview, 61
 predisposition *versus*, 70–72
 rational and revealed religion compared, 61–63
 rational religion and, 140
 Son of God and, 72n13, 126
 Walpole’s thesis and, 75
 wide duties and, 108
 evil, radical. *see also* change of heart; maxims,
 evil; self-deception; struggle against evil
 bad heart and, 145–146
 change of heart and, 88, 89–90, 104–105
 commentators on, 62n1, 69n10, 84–88,
 146n4
 concept of, 66–70
 concupiscentia and, 71n12
 condoning of, 147
 contemporary philosophers and, 62n1,
 68n10
 defined, 66, 84
 dispositions and, 102
 “egregious evil” and, 69n8
 the end justifies the means and, xvii–xvii
 everyday sins and, 62n1
 explaining, 66n4
 God’s commands and, 133
 Goethe on Kant and, xv
 heart and, 97
 human nature and, 86, 86n26, 87
 infinite punishment and, 132–133
 Kant’s “rigorism” and, 73–75
 literalism in religion and, 5
 moral and nonmoral incentives and,
 161
 Original Sin and, 61–63
 overview, 61
 passions and, 79–82
 practical freedom (free action) and, 63–66
 a priority and, 84–86
 “proof” of, 61, 84–88
 “radical” and, 62, 62n1
 rational and revealed religion compared,
 61–62, 87
 rational free action and, 66–70
 rational incentives and, 45, 89
 self-deception and, 91
 social evil *versus*, 66
 three degrees of, 82–84
 unsocial sociability and, 76–79, 85n25, 165
 Walpole’s thesis and, 75–76
 evil principle, 117, 117n2, 152
 evil reason, 68
 exceptions to moral rules, 73n15
 excommunication, 197
 “experiment,” 12n12
 experimental testing
 first experiment (duties and divine
 commands), 14–16
 overviews, 11–20, 12n12, 139–140
 rational religion and revealed doctrine
 compared and, 170
 second experiment (revelation), 11–13,
 13n13, 16–20, 17n16, 115, 116, 123,
 126, 150
 faith. *see* church and ecclesiastical faith
 faith, practical. *see also* disposition; hope;
 moral faith in God
 loyalty and, 37

- in *reality* of Son of God, 126–133, 129, 144, 146, 150
- faith and works, 178–182, 181n5
- “faith in search of understanding” tradition, 27n1
- falsity, 86
- fanaticism (*Schwärmerei*), 156, 199
- fear, 44, 197–198
- “fear of God,” 16
- feelings and attitudes, 12, 91–92, 94, 105, 106, 120, 153, 197. *see also* “felt need of reason”; inner feeling; love and other feelings; moral feelings
- felicity, 193, 194, 201
- fellowship, 185, 207
- feminist perspectives, 172n3
- fetish faith, 181, 213
- the *Fiat*, 149n6
- Fichte, Johann Gottlieb, xx, 20, 21, 141n2, 144n3, 146n4, 167n2
- fideism, 191
- field* (*Feld*), 13
- final end, 117
- final end of morality (*Endzweck*), 28, 50
- Firestone, Chris L., xiv, 17n16, 116n1
- Flikschuh, Katrin, 208
- foreign satisfaction (*fremde Genugtuung*), 180
- forgiveness. *see also* grace and salvation; mercy; receptivity
- condoning of evil *versus*, 147
 - grace and, 135
 - hope for, 138
 - Kierkegaard and, 146n4
 - legalism and, 146–147, 147n5
 - monetary sense of, 132, 133
 - morality and, 141n2, 141–142
 - proleptic, 137n14
 - punishment and, 137n13
 - rational and revealed religion compared, 142, 146, 149
 - religion and, 3
 - repentance and, 137n14
 - Son of God and, 135–138
 - as symbol, 144n3
 - vicarious atonement and, 146–148
- formula of autonomy, 204
- formula of humanity as end in itself, 204
- Formula of Realm of Ends, 166, 204
- formula of universal law, 204
- Formula of Universal Law, 84
- Formulas of Law of Nature, 84
- Das Forschen nach Licht und Recht in einem Schreiben an Herrn Moses Mendelssohn* ... (anonymous), 200n11
- fossil-fuel industry example, 30n5
- foundations of house, undermined, 75
- frailty (*Gebrechlichkeit, fragilitas*), 83, 98, 104, 105, 106, 107, 108. *see also* weakness of character
- Fränkel, David, 190
- Frederick the Great, 11
- free action. *see also* action and agency; causality; choice; disposition; duty; identity
- evil and, 63, 66–67, 68, 71, 75, 76
 - humanity and, 71
 - noumenal world and, 78
 - propensity to evil and, 63–66
 - rational, 5, 60, 66–70, 91
 - “two Worlds” and, 64, 78, 102n10
- free communication, 174, 175–176. *see also* conscience, freedom of
- freedom. *see also* autonomy; conscience, freedom of; identity; “individualistic”
- beauty and, 119
 - brutishness and savagery and, 71
 - desire and, 72
 - dispositions and, 96
 - ethical community and, 168–169
 - evil and, 63–66, 69
 - external, 81
 - faith and works and, 179
 - formal *versus* material, xx
 - God’s causality and, 66, 155–163, 156n10, 158n11, 181
 - grace and salvation and, 66, 156–158
 - habituation and, 107
 - highest good and, 34
 - inclinations and, 81, 93
 - metaphysical theories and, 64–65
 - moral incentive and, 153
 - moral law and, xx
 - as natural passion, 81
 - negative sense of, 75
 - noumenal world and, 64–66, 65n3, 78, 96, 97, 130n10, 160
 - possibility of, 64, 65, 156–157
 - propensity to evil and, 72
 - reform and, 183
 - “religious,” 213
 - of speech, 22, 199–200
 - supernaturalism and, 65n3
 - transcendental, 64, 65

- freedom (cont.)
 transcendental metaphysics and, 64n2
- French Revolution, 87
- Freud, Sigmund, 9n9
- Fricker, Miranda, 137n14
- Friedrich Wilhelm II, 11, 22, 23
- "from each according to his abilities ...," 110n15
- future conduct (second difficulty), 130–132,
 133, 134. *see also* moral improvement
- future life (immortality), 39, 46, 46n46, 191,
 200, 202
- future of religion, xx, xxi. *see also* reform
- future society, 110n15
- gallows and lust example, 76, 76n17
- German Enlightenment (*Aufklärung*), 10
- Germany, 7n8
- gloating, 79–80
- gluttony, 79
- God. *see also* anthropomorphism; forgiveness;
 God's commands; grace and salvation;
 kingdom of God, coming of; mercy;
 moral faith in God; service of God;
 "well-pleasing to God"
- agency of, 48
- causality and, 66, 155–163, 156n10,
 158n11, 181
- concept of, 4, 7, 10, 29, 118
- duties and, 142
- ethical community and, 168
- experience of, 124
- humanity and, 1n1
- idea of, 4, 15, 53, 119–120
- intuition of good disposition and, 129–130,
 130n9
- as judge, 143
- judgment of, 206
- knowledge and, 109
- love and, 135
- moral incentives and, 44
- moral requirements and, 180
- perfection and, 4n4, 15–16, 28, 120
- possibility of, 3
- power of, 34
- symbolically, 4n4, 28, 120, 173–174
- timelessness and, 128n8
- traditional concept of, 48, 50
- understanding and, 28
- will of, 28, 34, 206
- God's commands. *see also* experimental testing
 duties and, 14–15, 15–16, 176, 195, 210
 forgiveness and, 137
- God's cognition and, 4
- incentives and, 42
- moral duties and, 15, 210–211
- rational religion and, 12, 13, 14, 15–16,
 133–135
- religion and, 2, 139
- religious person and, 2, 3
- service of God and, 187
- sin and, 61–62, 133
- as symbols, 122, 125, 176
- "well-pleasing to God," 128, 128n8
- God's existence, faith in. *see* moral faith in God
- God's existence, proof of, 191, 201n13, 202n14
- Goethe, Johann Wolfgang von, xiv–xv, 87n26
- good, xviii, 37, 69n8, 74. *see also* highest good
- good, complete (*bonum consummatum*), 37
- good, natural, 38. *see also* happiness or well
 being
- "good" and "bad," 68n7, 73
- good common to all, 166
- good disposition, *see* change of heart
- good ends, xviii. *see also* the end justifies the
 means; ends of morality
- goodness of heart, 3n3, 97, 105, 141, 211. *see
 also* change of heart
- goodness of will, 37
- good predisposition, 153
- good principle. *see* "personified idea of the good
 principle"
- goods, 193
- Gorgias (Plato), 43n13
- grace and salvation. *see also* change of heart;
 forgiveness; God's causality; mercy;
 receptivity
- Augustinianism and, 140, 161n12, 161–163
- change of heart and, 110, 113
- Christian theology and, 140n1
- commentators on, 157n10, 158
- communion and, 185
- defined, 135
- empirical knowledge and, 111, 168
- experimental testing and, 140
- faith and, 52, 52n17
- forgiveness and, 141
- freedom and, 66, 156–158
- God's causality and, 155–163, 179
- historical faith and, 179–180
- hope for, 130
- justifying grace and, 140–142, 146, 151, 152
- Kant censored on, 21
- Lutheran Christianity and, 158, 178
- maxims and, 162n13, 162–163

- Mendelssohn on, 202
 miracles and, 121
 moral presuppositions of grace, 152–155
 “other” *versus* “better” human being and,
 154, 155
 overviews, 133–135, 139–140
 Pietism and, 9
 postulate of divine grace, 52n17
 prevenient grace, 151, 151n7, 152, 161
 rational religion and revealed doctrine
 compared, 139–140, 151–153,
 155–156, 157–158, 159, 161, 179
 reform and, 183, 212
 religious education and, 154–155
 resistance to, 163n15
 salvation by grace alone and, 151–152
 sanctifying grace, 140, 151, 152
 second experiment and, 19
 Son of God and, 148–151
 symbols and, 133, 161, 173n4, 202
 vicarious atonement and, 143–148
 “well-pleasing to God” and, 17, 146, 180
 gratitude, 3
 graven image, 6
 “Great men are almost always bad men.” (Lord
 Acton), 35n9
 greed, 80
 Greek ethics, 37. *see also* Epicureans; Stoicism
 Griswold, Charles L., 136n12
 grit, 56n18
 Grotius, 193n6
Groundwork of the Metaphysics of Morals
 (Kant), 39, 45, 74, 83, 91, 92, 98–102,
 100n9
 guilt (*Verschuldung*) (third difficulty). *see also*
 atonement, vicarious
 change of heart and, 131–133, 137
 collective, 166n2
 forgiveness and, 136, 136n12, 142, 147
 grace and, 140–141, 152
 irrational, 144n3, 144–145
 overview, 164
 punishment and, 135–136
 guiltlessness, 73
 “guise of the good,” 69n8
 Guyer, Paul, 37n10, 98n8

 habit or aptitude (*habitus*), 51, 57, 59, 60, 72,
 92, 96n5, 97, 105, 106–107
 habituation (*assuetudo*), 106–107, 108
 Hamann, Johann Georg, 6n7, 191
 happiness (*eudaimonia*). *see also* hope for
 happiness; worthiness to be happy
 duties and, 39, 99, 100
 as end of nature, 49–51
 eternal, 19
 examples and, 99
 freedom and, 81
 of individual *versus* world, 38
 injustice and, 43n13
 Judaism and, 202
 morality and, 40, 202
 of others, 39, 196n10
 pragmatic predisposition and, 71
 realm of nature and, 49–50
 reason and, 171
 religion and, 12
 of society, 49
 superiority over others and, 79
 happiness or well-being, 37
 Hare, John, xv, 2n2, 123–124, 137n13, 142,
 154n9, 159, 160–161, 189
 Harnack, Adolf von, 188n3
 Hasidism, 9
 Hasse, J. G., 11
 hatred, 69, 70n11, 92
 heart, goodness of, 3n3, 92. *see also* change of
 heart
 heart, perversity of, 62
 Heaven, 173
 hedonism, 70
 Hegel, Georg W. F., 141n2, 167n2
 Hell, 173
 helping others, 116
 Henry VIII example, 103
 Herder, Johann Gottfried, 191
 heresy, 14. *see also* anathema, Pelagianism and
 semi-Pelagianism
 Herman, Barbara, 145–146, 148,
 172n3
 “hermeneutics of suspicion,” xiv
 Hermes, J. T., 20–21
 “heroic” modern view, 74, 87
 heteronomy, 40, 44
 highest good (*summum bonum*). *see also ens*
 realissimum
 atonement and, 145
 ethical community and, 167
 as moral end, 37–39
 moral faith in God and, 33–39, 45–49, 50,
 52, 54, 57, 58, 59
 realm of nature and, 46n46
 second Critique and, 46
 symbols and, 143
 three Critiques and, 37n10

- highest good in the world (*Weltbeste*), 50
 Hillmer, G. F., 20–21, 21
 “Historical Inevitability” (Berlin), 59
 historical methods, 176
 historical phenomena, 6, 7–8, 13n13. *see also*
 church and ecclesiastical faith; neology;
 revelation
 historical truths, 4
 history, 211, 213
 holiness, 72n13, 106, 107, 116–117. *see also*
 moral progress
 Holocaust, 70
 holy will, 63, 72n13, 116–117, 143
 honest merchant example, 98–99
 honor, 36, 79n23, 80, 81, 82, 169, 170, 170,
 194–195
 hope
 acting from duty and, 109
 belief and conviction *versus*, 197–198
 change of heart and, 17, 109, 111, 112, 113,
 128, 129, 144
 for forgiveness, 138, 144
 for future contingent events, 51–52
 for grace and salvation, 130
 Kant’s, 210, 213–215
 for mercy, 138n16
 moral argument and, 40, 210
 “Ought implies Can” and, 127
 practical faith and, 51, 51n16, 52n17, 138,
 138n15
 rational and revealed religion and,
 115
 rational faith and, 44
 rational religion and, 138
 for redemption, 115
 Richard II and, 59
 for satisfaction of needs, 43
 for transition of religion, 184
 to well-pleasing to God, 152
 hope for happiness (rational, moral)
 change of heart and, 17
 moral disposition and, 41
 moral incentive and, 40, 42,
 44
 second Critique and, 45
 worthiness and, 42, 44, 46
 hopelessness, 59, 131
 “how great God is,” 15
 Hudson, Hud, 64n2
 humanity, predisposition to, 71–72,
 77, 79
 “humanity in its full moral perfection”
 (formulation of ideal of humanity),
 116–117
 human nature, 169. *see also* evil, propensity to;
 evil, radical; frailty; sensuous nature;
 unsocial sociability
 Hume, David, xx, 6, 7, 15, 27n1, 79, 94,
 103, 213
 humiliation, 50
 hypocrisy, 86, 207
 hypotheses, 53
 hypotyposis, 118, 120, 124, 125

 iconoclasm, 173n4
Idea for a Universal History (Kant), 86, 171
 “ideal,” 116
 ideal ethical community, 16, 166, 168
 “ideal of holiness” (formulation of ideal of
 humanity), 116–117. *see also* holiness
 “ideal of humanity well-pleasing to God”
 (formulation of ideal of humanity). *see*
 “well-pleasing to God”
 ideal of pure reason, 121
 ideals, 38, 118, 187
 idea of God, 4, 15, 53, 119–120
 ideas, 118
 ideas, aesthetic, 6
 ideas of reason, 4, 5, 15, 124–125. *see also* God
 an other ideas of reason
 identity (as free, self-governing, rational agent),
 82–83, 91, 105, 112, 113, 113n16
 idiosyncrasy, 208
 idolatry (*avoda zera*, *shirk*), 7–8, 122, 122n4,
 123, 124, 173, 173n4, 213
 idols, 121
 ill (*Übel*), 66, 134
 ills of life, 144, 148
 illusions, 59, 112
 imagination, 1
 immaturity, 63
 Immediate Commission of Investigation, 11
 immortality (future life), 39, 46, 46n46, 191,
 200, 202
 imperfection, 113, 128n8, 168
 imprudence, 70n11
 impurity (*Unlauterkeit*, *impuritas*, *improbitas*),
 83–84, 108
 imputability
 conscience and, 205–206
 deeds *versus* inclinations and, 72
 of evil, 62–63, 66, 67, 68, 85
 free agency and, 63, 75

- God's causality and, 160
 of good and evil, 154–155, 162
 grace and, 154–155, 162
 individual responsibility and, 166n2
 moral predisposition and, 71
 natural predispositions and, 70
 propensities and, 72
 inaction, 163, 163n15
 Incarnation, 150, 183
 incentives, moral and nonmoral. *see also*
 happiness and other nonmoral
 incentives; hope for happiness *and other*
 moral incentives; Walpole's thesis
 acting from/in conformance with duty and,
 74, 96–97, 99–101, 100n9, 108–109
 bad heart and, 143
 empirical motivation and, 173
 evil and, 63, 68, 69, 75–76, 159, 161
 good predisposition and, 153
 passion and, 80
 self-deception and, 83–84, 91
 incentives (*Triebfeder*) and motives. *see also*
 happiness and other incentives;
 incentives, moral and nonmoral; moral
 incentive; self-deception
 acting from, 94–98
 ambiguity of, 96n4
 change of heart and, 90–91, 94–98
 as characterization or way of acting, 95, 96
 emotions and, 174
 empirical, 173
 goodness of heart and, 105
 good predisposition and, 153
 habituation and, 108
 justice and, 166
 nonmoral, 74
 passions and, 80
 practical freedom and, 65
 rational, 68
 self-deception and, 80
 inclinations (*Neigung*). *see also* passions
 change of heart and, 105
 concupiscence and, 72
 evil and, 63, 67, 70, 80, 84
 immediate, 74
 possibility of, 71–72
 reason and, 82
 unsociable sociability and, 77
 virtue and, 93
 “incorporation thesis,” 93
 incredulity, 60
 indebtedness (*Verschuldung*), 131
 indifference, moral, 59, 73n14, 73–74, 169,
 192, 212
 individualism, 166–167, 167n2, 189, 208
 “individualistic,” 166n2
 inducement to do things to make us believe, 30
 inextirpable (*unvertilgbar*), 63, 85, 85n24,
 102–103, 105, 106
 influences, xx
 ingratitude, 69, 70n11, 71, 79, 86
 injustice, 43n13, 66, 162n13
 innateness, 72n13, 81, 84, 85
 innateness and Original Sin, 62–63
 inner feeling, 177–178
 inner revelation, 174
 inner state, 186
 insanity, 63
 Insole, Christopher, xv, 157n10
 instincts, 70, 72, 79
 intelligible *versus* sensible concepts, 156
 interdependence, 193
 intolerance, 213
 intuition, 5, 117–118
 irrationality, xx, 6, 7n8, 25–26, 191. *see also*
 superstition
 Islam, 6, 7, 11, 142, 173, 177, 212. *See also*
 monotheism
 Jacobi, F. H., 11n11, 191
 Jacobs, Nathan, xiv, 17n16, 116n1
 James, William, xx, 58
 Jansenism, 162n13
 Jaspers, Karl, 70
Jerusalem: Or, On Religious Power and
 Judaism (Mendelssohn), 191–192, 203
 Jesus Christ
 change of heart and, 111–112
 existence of, 128
 Lutheranism and, 140
 masculine noun and, 89n1
 rational religion and revealed doctrine
 compared, 115–116
 resurrection of, 150
 unity with, 124
 “well-pleasing to God” and, 180
 joy in another's suffering (*Schadenfreude*), 71
 Judaism and Jews. *see also* monotheism;
 toleration
 Christianity compared, 191, 203, 203n15
 compensation for rabbis and, 196
 conception of God and, 29, 177
 end of the world and, 195n8
 enlightenment and, 177

- Judaism and Jews (cont.)
 evil symbols and, 7n8
 freedom of conscience and, 192, 200
 Hasidism, 11
 legalism and, 192n4, 201, 202–203
 Mendelssohn and, 191–192, 200–201, 203
 mercy and (Yom Kippur), 142
 Pentecost and, 110n14
 Prussia and, 192n5
 rational religion and, 192n4, 201–203
 respect for, 212
 scripture and, 177
 separateness and, 203
 the sublime and, 6, 173
 judgment about others, 112
 judgment and conscience, 204n17, 204–205
 judgments, ethical community and, 166
 justice, 16, 38, 60, 147n5, 148, 166, 194
 “justification,” 137–138
 justifying grace, 140–143, 146, 151, 152
- Kaddish* (Symphony no. 3) (L. Bernstein), 186n2
- Kahn, Samuel, 34n8, 75n16, 204n17
- Kail, P. J. E., 6n7
- Kang, Y., 5n6
- Kant, Immanuel. *see also* Mendelssohn, Moses
 censored, 21, 185
 influence of, 1–2
 interpretations of, xiv–xx
 life and religion of, xiv, xix, 9–14, 110, 185, 191, 210
 Mendelssohn and, 111n11, 191, 201–209, 202n14, 203, 203n15, 208, 212
 “Kant’s Rejection of Divine Revelation” (Savage), 13n13
- Kierkegaard, Søren, xx, 12–13, 141n2, 146n4
- “kind of person,” 97
- King, Jr., Martin Luther, 7n8, 60
- kingdom of God, coming of, xxi, 181, 182–184, 211
- Kiesewetter, J. G. K. C., 21
- knowledge (*Wissen*), 17, 30, 41–42, 44, 52–53, 56, 123–124. *see also* empirical (psychological) evidence and experience; physician helping patient example
- Knutzen, Martin, 10
- Kristallnacht, 7n8
- “laborer in the moral vineyard,” 48–49
- laity-clergy distinction, 182–183
- Lamentabili sane exitu* (Pius X), 188n3
- Lamott, Ann, 56
- Lavater Johann Kaspar, 191, 200n11
- law of autonomy, 172
- law of freedom, 39
- laws. *see also* moral law
 churches and, 196
 ethical community and, 166–168, 169, 170
 Jewish (Mosaic Law), 200–201, 201n13, 202, 203
 religion and, 192n4
 revealed, 200–201, 207
- laws of nature, 39, 120, 121n3. *see also* causality
- laws of understanding, 121
- “learned public,” 23
- Lectures on the Philosophical Doctrine of Religion* (Kant), 15, 41n12, 42, 44
- Lee, Robert E., 7n8
- legality; legalism and legal metaphors. *see also* advocate; debts; laws
 conscience and, 205
 enlightenment *versus*, 121
 forgiveness and, 146–147, 147n5
 guilt and, 133
 imputability and, 155
 mercy and, 126
 morality and, 94, 101–104, 107, 109, 167
 moral reason and, 92
 struggle against evil and, 115
- The Legend of the Grand Inquisitor* (Dostoevsky), xvi
- Leibniz, G. W., 29
- Lessing, G. E., 7, 10–11, 111n11, 191, 201n13, 212
- liberalism, 21, 188n3
- lies, 5, 69, 70n11, 79. *see also* self-deception
- “liquidating the Just Society” (McBride), 66n4
- literalism, 5, 9, 14, 14, 134, 151–152, 173, 211. *see also* anthropomorphism; atheism; symbols, religious
- Locke, John, xx, 52, 194
- Longuenesse, Béatrice, 172n3
- “lost” causes, 7n8, 36, 38, 59
- Louden, Robert, 67n5
- love, 3n3, 16, 21, 80, 82, 92, 106, 119, 135, 142
- love and respect in balance, 141, 143
- love of life, 99
- “Love your enemies ..., ” 177
- loyalty, 35–37, 59, 210
- lust and gallows example, 76, 76n17
- Luther, Martin, 185n1
- Lutheranism. *see also* Pietism

- “alter Mensch” and “neuer Mensch” and,
 89–90
 censorship and, 11, 20–21
 criticism of, 21
 experimental testing and, 19
 grace and, 158, 178
 Jesus Christ and, 140
 Original Sin and, 86
 other communities and, 212
 Prussia and, 194n7
 reform and, 188
 theology and, 10
- Madhyamika Buddhism, 3n3
 malice, 69, 70n11, 86
 Mandela, Nelson, 137n14
 mania, 79
 Mariña, Jacqueline, 37n10, 140n1, 161n12
 Maritain, Jacques, 27n1
 market exchanges, 193–194
 Marx, Karl, 110n15
 Marxism, 59, 142, 167n2
 material principles, 91
 Matušík, Martin Beck, 68n7
 maxim of holiness, 105
 maxim of security (argumentum a tuto), 207
 maxims
 acting from duty and, 107
 change of heart and, 89, 106, 110–111
 empirical experience and, 66
 fundamental, 97, 110
 grace and, 162n13, 162–163
 of holiness, 105
 passions and, 79
 “rigorism” and, 73
 self-deception and, 74
 Stoic, 154
 maxims, evil, 67, 69, 72–75, 76, 78, 83, 93,
 137, 154
 McBride, William, 66n4
Meditations (Descartes), 57
 Mendelssohn, Moses
 on church, 197n10
 on churches, 192–197, 194n7
 on different ecclesiastical communities, 212
 freedom of conscience and, 206
 freedom of conscience and evidentialism and,
 197–200
 Judaism and, 191–192, 200–201
 Kant and, 11n11, 191, 201–209, 202n14,
 203, 203n15, 208, 212
 life and works of, 11n11, 190–191
 on natural right, 192–197, 198–199,
 201–202, 203
 “new and harsh” doctrines and, 196n9,
 196–198
 overview, 186–187
 on Prussian society, 192n5
 rational and revealed religion compared,
 201–202, 202n4
 on scripture and tradition, 201n13
 separation of church and state and, 194n7
 on state, 194n7, 194–195, 196, 197, 199–200
 mercy, 136, 137n14, 138n16, 142, 143, 147,
 148, 180, 187. *see also* forgiveness
 “mere (*bloß*) reason,” 15, 24
 Merritt, Melissa, 96n5
 metaphysics, 66
Metaphysics of Morals (Kant), 205
 Mexico, border wall with, 7n8
 Michaelis, J.D., 176–177
 Michalson, Gordon, xii, xv, 28, 29n3, 29–30,
 38, 50, 85n24, 87n26, 107n24, 128n8
 military force, 171
 Mill, John Stuart, 196n10
 “minimum of theology,” 52
 miracles, 121, 121n3, 175
 miracles, “moral,” 160
 modernity, 9, 20, 87, 188, 188n3, 214–215.
 see also contemporary culture
 money, 80, 82
 Montaigne, Michel de, 77
 Moors, Martin, 13n14, 116n1
 moral argument. *see* moral faith in God
 moral cooperation (*concursum morale*), 157
 moral discipline, 61, 103. *see also* discipline,
 culture of
 “moral evidence,” 29, 29n4
 moral faith in God. *see also* faith, practical
 antievindentialism and, 56n18, 56–61
 assent and, 31–33, 51–56
 in *Critique of the Power of Judgment*, 34,
 49–50
 in *Critique of Practical Reason*, 34, 45–49
 in *Critique of Pure Reason*, 34, 39–45
 duty and, 15, 241n12
 failure to accept, 47
 highest good and, 33–39, 45–49, 50, 52, 54,
 57, 58, 59
 hope and, 40, 210
 as incentive, 41n12
 knowledge *versus*, 5, 15
 “moral evidence” and, 29, 29n4
 nonbelievers and, 35

- moral faith in God (cont.)
 overviews, 1, 2, 10, 15, 28–29, 35, 210
 philosophical arguments *versus* believers
 and, 27n1, 27–29
 practical grounds of assent and, 29–36
 practical *versus* theoretical reason and,
 47–48
 pure reason and, 44–49
 three Critiques and, 34–35
 three Critiques compared, 34–35, 45
- moral feelings, 71, 82
- moral good, 37, 38. *see also* good; goodness of
 will; virtue
- moral improvement. *see also* struggle against
 evil
 change of heart and, 93, 143
 ethical community and, 18, 189, 211
 hope and, 131, 138
 irrational guilt and, 145
 moral incentive and, 90–91
 rational and revealed religion compared, 156
- moral incentive. *see also* incentives, moral and
 nonmoral
 change of heart and, 89–98, 89–98, 106,
 108–109, 110–111
 empirical evidence and, 40–42, 41nn11, 12,
 43, 95
 faith and, 41n12
 first Critique and, 39–40, 41n12
 good heart and, 92
 identity and, 82–83, 91
 Kant's "rigorism" and, 73
 moral argument and, 39–46
 moral law and, 39, 68, 69, 71, 73
 psychology and, 40–42
 purity of, 90–91
 rationality and, 42–45
 second Critique and, 45–49, 92
 "sufficiency" of, 97, 97n6, 106, 109
 three Critiques compared, 45
 weak, 41, 41n11
- moral intermediates, 73, 73n14
- morality. *see also* ethical community; legalism
 and legal metaphors; legality; moral
 faith in God; moral incentive; moral law;
 reform; secularism
 Abraham and Isaac and, 14
 beauty and, 122
 Christianity and, 12, 14, 14n15
 church and, 202, 214
 end of, 33–34
 evil and, 140
 forgiveness and, 142
 happiness and, 40
 history of, 171
 legality and, 94, 101–104, 107, 109, 167
 modern world and, xii
 pragmatic reason compared, 39
 rational and revealed religion compared,
 17, 182
 rational religion *versus*, 139
 reason and, 25, 68
 religion and, 2–3, 6, 10, 11–12, 14, 187–188,
 210–211
 scripture and, 214
 social origins of, 172
 tradition and, 202
- moral law. *see also* obedience
 apriority and, 90
 conformity to, 37
 emotions and, 172
 first Critique and, 39
 freedom and, xx
 God's will *versus*, 1
 happiness and, 202
 highest good and, 38
 incentives to disobey, 68
 letter *versus* spirit of, 94
 moral incentive and, 39, 68, 69, 71, 73
 respect for, 92
 revelation and, 18
 for sake of, 90, 94
 "moral miracle," 160
- moral progress. *see also* enlightenment
 change of heart and, 111, 113–114, 129
 duty to act from duty and, 107
 empirical and practical evidence and, 87
 holiness and, 129, 130
 modernity and, 20
 propensity to evil and, 87, 104–105
 rational and revealed religion and, 180
 self-deception and, 130
 time and, 113–114
- moral religion. *see* rational (moral) religion
- Moral Religion, 52–53
- moral weakness, 83
- Moravianism, 153
- Morgan, Seiriol, 78n21, 85n25
- Morgenstund* (Mendelssohn), 202n14
- Mörschel, Daniel Ernst, 200n11
- Morton, Jennifer M., 56n18
- Muchnik, Pablo, 85n25
- Murder in the Cathedral* (Eliot), 113n16
- music, xix

- Muslims, xix
 mysticism, 124n6, 153
- Nathan the Wise* (Lessing), 191
 natural good, 37, 38
Natural History of Religion (Hume), 6
 naturalism, 18, 19
 natural law, 192–193
 natural passions, 80–82
 natural religion, 24
 natural rights, 192–197, 198–199, 201–202, 203
 natural theology lectures (Kant), 157
 natural world, 214
 nature and realm of nature, 1, 46n46, 49–50, 77. *see also* realm of nature; science
 necessitation, 95
 neology, 10, 11
 Neuhausser, Frederick, 77n18
 Nicolovius of Königsberg, 21
Nicomachean Ethics (Aristotle), 31, 49, 72
Nihil appetimus, nisi sub ratione boni; nihil aversamur, nisi sub ratione mali (“We desire nothing except under the reason of the good; we avoid nothing except under the reason of the bad”), 68n7
- nihilism, xv
 Norris, Henry, 103
 “nothing,” four kinds of, 65
 noumenal world. *see also* the supersensible;
 transcendent metaphysics
 change of heart and, 112, 113
 evil and, 78, 97
 freedom and, 64–66, 65n3, 78, 96, 97, 130n10, 160
 responsibility and, 160
 self-opacity and, 102n10
 “two perspectives” problem and, 128n8
 virtue and, 93, 94, 105–106
- Obama, Barack, 7n8, 60
 obedience, 16, 50, 166
 “objectively considered,” 2
 “old man” and “new man” (“alter Mensch” und “neuer Mensch”), 89, 89n1, 105, 125, 134, 137n13, 144, 146, 146–147, 148, 148–149
On the Civil Advancement of the Jews (Dohm), 196–197
 One Divinity, 3n3
On the Common Saying (Kant), 21
 ontological argument, 27n1
 ontotheology, 119–120
 opinion (*Meinung*), 53
 opinion of others, 80, 172
 oppression, 212
Opus Postumum (Kant), 34n8, 37n10
 Original Sin, 19, 21, 61–63, 71n12, 86, 87n26, 140
 “other” human beings and “better” human beings, 153–154
 other people, 39, 86, 112, 193, 196n10, 211, 213. *see also* relations with others;
 unsocial sociability
 comparison of self with, 71, 76, 78, 78n23, 80
 “Ought implies Can,” 41, 75, 75n16, 103, 127, 158–159, 159, 160–161
 overconfidence, 59, 60
- pagan ideas, 6
 Palmquist, Stephen R., xiv, 2n2, 28n4, 85n25, 116n1, 117n2, 128n8
 Papish, Laura, 66n4, 67n6, 85n25, 109n12, 113n17
 parables, 6
 partisanship, xx
 Pascal, Blaise, xx, 30, 30n5
 passions, 79–82
 Pasternack, Lawrence
 on divine justice, 135, 135n11
 on experimental testing, 12n12
 on forgiveness, 147
 on grace, 158, 158n11
 Kant’s Christology and, 116n1
 on Kant’s “guile,” 22
 on philosophical *versus* theological doctrines, 25n20
 on pure rationalism, 18n17
 on revelations as “anathema,” 14n15
 on symbolism, 5n6
 “well-pleasing to God” and, 128n7
 Paul, Saint, xviii, 75–76, 94, 151
 Paul, Sarah K., 56n18
 peacock of Juno, 6
 Pelagianism and semi-Pelagianism, 62, 161, 178, 179. *see also* heresy
 Pentecostals, 110, 110n14
 the people (*Volk*), 167–168
 “people of God,” 18, 211
 perfectibility, 76–77
 perfection, moral, 46n46
 perfection of humanity, 116–117
 perfect will, 127

- persistence, 56n18
 personality, 71
 "personified idea of the good principle"
 (formulation of ideal of humanity),
 72n13, 116–117, 117n2
 persuasion (*Überredung*), 41n11, 52
 perversity, 62, 83, 84
 petitions, 193, 194
Phantast (visionary), 42, 43
 Philip J. Rossi, 5n6
 philosophical faculty at University of
 Königsberg, 21, 23–24, 25, 25n20
 philosophy of history, 77
The Philosophy of Liberty (Royce), 35–37
 physician helping patient example, 32–33, 34,
 53, 53–54
 physicians, 23
Pia desideria (Spener), 153–154
 Pietism, 9, 9–10, 110, 110nn13, 15, 111, 113,
 114. *see also* Spener, Philip Jacob and his
 problem
 Pius X, 188n3
 Plato, 43n13
 "playing the victim," 147n5
 pleasure, 79, 86
 Pluhar, Werner, 158n11
 pluralism, 183n6, 212
Plus de Girondins (Blanc), 110n15
 political beliefs, 207
 political history, 176, 182. *see also* states,
 political
 politics, 213
 politics, contemporary, xvii–xix, xxi, 145, 184,
 207, 215
 polytheism, 3, 3n3
 popular moral philosophy, 91
 popular religion, 6
 possibility. *see also* "Ought implies Can"
 of actions, 47
 of an inclination, 71–72
 assent and, 57
 of ends, 60
 of forgiveness, 141
 of freedom, 64, 65, 156–157
 of God, 3, 46–47, 52, 119–120
 of highest good, 33–34, 39, 46–47, 54, 55, 57
 of highest good and, 33–34, 39
 logical *versus* real, 4, 5, 6, 33–34
 of motivations, 75n16
 practical grounds and, 31, 34
 of religion, 120
 of revelation, 19
 symbols and, 4n4, 5, 34
 of "well-pleasing to God," 125, 128
 postulate of divine grace, 52n17
 postulate of pure practical reason, 47–48
 postulates, 52n17
 power, 77, 78n23, 80, 82
 "Power corrupts; absolute power corrupts
 absolutely" (Lord Acton), 35n9
 practical reason. *see also* moral faith in God
 assent and, 46
 highest good and, 47–48
 moral incentive and, 39, 40–41
 possibilities and, 31, 34
 theistic arguments and, 28
 practical standpoint
 experience of God and, 5
 God's existence and, 15
 symbols and, 4
 pragmatic belief, 32, 33
 pragmatic Belief, 34, 53, 54, 57
 pragmatic predisposition, 71
 pragmatic reason. *see* prudence
 praise, 74, 172
 prayer, 185, 186n2
 predestination, 66, 162, 162n13
 predisposition (*Anlage*), 70–71, 74, 77. *see also*
 evil, propensity to; instincts *and other*
 predispositions
 presentations (*Darstellung*), 118, 119, 121,
 124, 125
 preservation of species (sexuality), 70, 145,
 172, 184
 prevenient grace, 140, 151, 152, 161–162
 Prichard, H. A., 95n3, 128n8
 priestcraft (*Pfaffentum*), xxi, 121, 171, 175,
 181, 186–187, 190, 196, 213. *see also*
 laity-clergy distinction
 "problems," 47, 47n15
 progress in history, xix, 211
 propensities. *see* animal nature; evil,
 propensity to
 property rights, 171–172, 193, 196
 propositional attitudes, 50, 51, 52
 propriety (*Sittsamkeit*), 172
Proslogion (Maritain), 27n1
 Protestantism, 185n1, 188n3, 212
Protocols of the Elders of Zion, 7n8
 "prototype of moral disposition in its entire
 purity" (formulation of ideal of
 humanity), 72n13, 116–117, 150. *see*
 also disposition
 prudence, 39, 70, 74, 80, 98–99

- prudential knowledge, 44
 Prussia, xv, 10, 11, 20–25, 185, 192nn4, 5, 194n7. *see also* censorship; Lutheranism
 Psalm 59, 176–177
 psychotherapy, 145, 146
 public good, 199
 public legislation, 167
 public well-being, 194
 public worship, 121
 Pufendorf, Samuel, 192–193, 193n6
 punishment and penalties
 blame and guilt and, 135–136
 churches and, 196, 197
 convictions and, 197–198
 divine justice and, 134
 duties and, 196n10
 fear of, 44, 197–198
 forgiveness and, 137n13
 freedom of conscience and, 206, 207, 208
 guilt and, 132–133, 137n13
 moral incentives and, 41–42
 self-regarding conduct and, 196n10
 as symbolic, 146–147
 vicarious atonement and, 146–147
 willing one's own, 134
 pure concepts, 5, 118
 pure disposition, 93, 109, 111, 126–131, 138n16, 143–145, 147–149, 151–152
 pure practical reason, 39, 47–48
 pure rationalism, 18n17, 18–19
 pure reason, 45–49
 purity of ecclesiastic faith, 123, 168
 purity of ethical community, 168
 purity of moral incentive, 90–91

 Quakers, xix, 197
 questioning, 59
 Quinn, Philip, 29

 rank, 82
 rational and revealed religion compared
 (concentric circles image). *see also*
 ethical community; grace and salvation
 and other doctrines; morality;
 revelation; scripture; service of God;
 symbols, religious
 being convinced of a higher being and, 174
 Belief and, 111, 152
 cognition of the divine and, 173, 174
 commentators on, 13n3, 14n5, 18n17, 116n1
 enlightenment and, 177, 184, 187
 Lessing and, 7
 mutual dependency and, 17, 163, 173, 187, 211, 213
 one true religion and, 2, 201n12
 overviews, 11–14, 17, 18–20, 24–25, 115–116, 139–140, 150–151, 163, 211, 213
 reason and, 178–182
 transition between rational and revealed, 178–182, 183
 rational nature as end in itself, 15
 rational (moral) religion, xiii–xiv, 1–2, 19, 139, 201–211. *see also* ethical community; God's commands; rational and revealed religion compared; reform; symbols, religious; unity and universality
 rational self-interest, 70n11
 rational subjective and practical grounds, 33
 rational volition, 84
 Ray, Matthew Alun, xv
 Raz, Joseph, 69n8
 realm of ends, 3n3, 16. *see also* community of rational agents
 realm of nature, 1, 46n46, 49–50, 77. *see also* science
 reason, rationality and rationalism. *see also*
 conviction; enlightenment; moral faith in God; morality; rational (moral) religion
 belief and, 201–202, 210
 belief in God and, 29
 church and, 25–26, 181
 cognition of God and, 5
 contemporary defiance of, 213
 dogmatism and, 25–26
 emotions and, 3n3
 Enlightenment and, 1, 213
 “felt need of,” 54
 happiness and, 171
 identity and, 82–83
 incentives and, 68
 judging itself, 205
 mere (*bloß*) reason, 15, 24
 moral incentive and, 42–45
 psychological incentives and, 43
 pure, 18n17, 18–19
 rational and revealed religion compared, 178–182
 revelation and, 175, 214
 self-love and, 77
 superstition and, 120–121
 time and, 97
 reasons

- reasons (cont.)
 to assent, 2–3
 to believe, xx, 29–30
 causes and, xx, 96n4
 as characterization or way of acting, 95, 96
 empirical, 93
 evil and, 66–67
 free action and, 66
 instrumental or prudential, 30, 68
 “kind of person” *versus*, 97
- Reath, Andrews, 48–49
- rebirth, 89, 105, 110
- receptivity, 135, 137, 138, 142, 143, 144,
 151n7, 162–163, 180
- reconcilability (*Versöhnlichkeit*, *placabilitas*),
 135, 135n11, 142
- redeemer (*Erlöser*), 134
- redemption, 115, 132, 203n15. *see also* grace
 and salvation; Son of God
- redemption of world, 203n15
- reductionism, 78
- reform. *see also* moral improvement
 of Christianity, 181–184, 203, 211, 212
 contemporary religion and, 184, 213
 ethical community and, 171, 189–190, 211
 freedom of conscience and, 207–208, 212
 historical necessity and, 189
 Mendelssohn and, 192, 208, 212
 rational and revealed religion and, 17, 19–20,
 182, 187–188, 211, 212, 213
 of service of God, 183, 187–189
 symbols, religious and, 174, 182, 188–189,
 211, 214
 unity/universality and, 182, 187, 203, 211,
 212, 213
- Reimarus, Hermann, 10, 10n10, 201n13
- Reinhold, Karl Leonhard, 34n8, 65n3
- relations with others, 166, 194, 195. *see also* ethical
 community; unsocial sociability
- religion, 1–3, 6–11, 195, 211, 213–214. *see also*
 church and ecclesiastical faith;
 enlightenment; evil, radical; revelation;
 states, political; symbols, religious
- Religion within the Boundaries of Mere Reason*
 (Kant), xiii, xiii–xxi, 1, 209
 overviews, 1–2, 164
 publication of, 20–26
- religious attitude, 52
- religious communities, xxi, 110n15
- “religious fetishism,” 28
- religious instruction, 154–155
- religious people, xix, 54–55
- religious practice. *see also* service of God
 arguments for belief in God and, 27
- repentance, 136, 137nn13, 14, 142
- republican constitution, 183n6, 183–184
- resentment, 92, 137n14, 196n10
- resignation, 55, 60
- respect, 3n3, 16, 82, 92, 141, 143, 172,
 211, 212
- responsibility, 5, 159, 160
- resurrection of Christ, 150
- revelation. *see also* church and ecclesiastical
 faith; rational and revealed religion
 compared; symbols, religious
 evil and, 73–75
 experimental testing and, 11–13, 13n13, 15,
 16–20, 17n16
 hope for redemption and, 115
 Judaism and, 200
 laws and, 200–201, 207
 moral concepts and, 17
 moral law and, 18
 need for, 170–171
 nonempirical, 183
 rational and revealed religion compared,
 174–175, 181–182
 rational religion and, 7, 18–19, 24
 reason and, 175, 214
 scripture and, 18
 service of God and, 16–17
 sublimity and, 172–174
- revivalist movements, 9
- “revolution,” 110, 110n15
- rewards, 196, 197–198, 206, 207, 208
- Richard II* (Shakespeare), 59
- right and wrong action, 204–205, 205n18
- rights
 to believe according to reason, 201–202
 churches and, 196, 197
 God’s, 194–195
 individual, 183
 natural, 192–197, 198–199, 201–202, 203
 other people’s, 197n10
 perfect and imperfect, 193n6, 193–194,
 195–196
 to reason, 198–199
 theories of, 9, 23, 166n2, 183, 193
 “rigorism,” Kant’s, 73nn, 14–16, 73–75
- Ritschl, Albrecht, 188n3
- rituals, 185–186
- Roman scripture, 177
- Rosenzweig, Franz, 203
- Ross, W. D., 5n6, 95n3, 128n8

- Rousseau, Jean-Jacques, 76–77, 77n18, 80
 Royce, Josiah, 35–37, 38, 59, 210
- Sachs, Albie, 142
 Sala, G. B., 116n1
 salvation. *see* grace and salvation
 sameness, 118, 119, 120, 124, 125, 126
 Sartre, Jean-Paul, 56
 satisfaction, 43, 72, 134, 137, 152, 178, 179, 180
 Savage, Denis, 13n13
 savagery, 71, 81
 Schapiro, Tamar, 73n16
 schemata, 118, 125
 Schleiermacher, Friedrich Daniel Ernst, 65n3, 188n3
 scholarship, recent, xiv–xv
 scholasticism, 48, 68n7
 Schopenhauer, Arthur, 45, 94
 Schott, Robin May, 172n3
 Schrift, A., 67n5
 Schulz, Franz Albert, 9
 science, 4, 7, 8–9, 86, 113, 176, 188, 201n13, 211, 213, 214
 scientism, 8–9
 Scotus, John Duns, 29
 scoundrel (*Bösewicht*), 42, 43
 scriptural scholarship, 177
 scripture. *see also* Bible; revelation
 churches and, 169, 190
 enlightenment and, 212–213
 as evidence, 201n13
 interpretation of, 174–178
 moral content and, 214
 rational and revealed religion compared, 19, 174–175, 177
 reform and, 183
 revelation and, 18, 214
 symbolism and, 6, 7, 188, 201n13
 secularism. *see also* states, political
 forgiveness and, 142
 highest good and, 48–49
 Kant and, xiii, 1–2, 182, 183, 188, 190, 210, 211
 “lost” causes and, 36–37
 modern, 184, 213, 215
 religion and, 215
 symbolism and, 4
 self-acceptance, 137, 143, 144n3, 146, 146n4
 self-caused cause, 95
 self-conceit (*Eigendünkel*, *arrogantia*), 71, 77
 self-constraint, 91, 100n9, 101
 self-deception. *see also* superstition
 absence of, 91
 acting from duty and, 98
 as always a vice, 108
 Belief and, 60
 certainty and, 205
 change of heart and, 105, 109n12, 130–131
 counterfeit service and, 185
 evil and, 67, 67n5, 69, 74, 80, 82–84, 91
 freedom and, 66
 impurity and, 83–84
 maxims and, 74
 moral progress and, 129, 130
 promised advantage and, 41n11
 psychological incentives and, 41nn, 11, 12, 41–42
 self-examination, 102, 103
 self-flattery, 108, 109
 self-forgiveness, 143
 self-honesty, 66, 112–113
 selfhood, 146n4
 self-indulgence, 204
 self-indulgent hedonist, 70
 self-interest, 70, 70n11. *see also* “Every man has his price.”; honest merchant example
 selfishness, 100
 self-knowledge, 102, 102n10, 112
 self-love, 70, 71, 77, 84, 94, 108, 112
 self-love (*Eigenliebe*, *philautia*), 77
 self-love (*Selbstliebe*, *solipsismus*), 70–71
 self-opacity, 74, 102, 102n10
 self-pity, 147, 147n5, 148
 self-preference, 204
 self-preservation, 70, 76, 76n17, 79
 self-regarding conduct, 196n10
 self-reproach, 50
 self-salvation, 178
 self-seeking, 98–99, 100
 sense and significance (*Sinn und Bedeutung*), 4, 4n4, 5
 sensibility, 117–118, 124
 sensible impulses, 44, 63–64
 sensible *versus* intelligible concepts, 156
 “sensitive but free” (*arbitrium sensitivum sed liberum*), 65
 sensuous nature, 67, 68
 sensuousness, 172–173
 sentiment, 3n3
 separation of church and state, 194n7. *see also* states, political
 service of God. *see also* creeds and other practices

- service of God (cont.)
 counterfeit (*Afterdienst*), 181, 185, 185n1,
 187–190, 202–203, 213
 critiqued, 187–190
 ethical communities and, 169
 faithful, 119
 fetish faith and, 181
 God's need and, 194–195
 Kant and, 185
 rational and revealed religion compared,
 187–189
 reform and, 183, 187–189
 revealed laws of, 168
 revelation and, 16–17
 rituals and, 185–186
 true, 170, 212–213
- servility, 79
- Setiya, Kieran, 69n8
- seven is not prime bribe, 41n11, 52
- sexuality (preservation of species), 70, 145,
 172, 184
- sexual passion, 76n17, 81–82
- Shakespeare, William, 59
- shame, 43n13
- “shrunk God,” 29, 50
- shunning, 197
- The Sickness unto Death* (Kierkegaard), 146n4
- Sidgwick, Henry, 65n3, 144n3
- Silber, John, 68n7, 128n8, 141n2
- sin, 61–62, 62n1, 75–76, 131–133, 140, 212.
see also Original Sin
- Sinclair, Upton, 30n5
- 6000 (Jewish year), 195
- skills and culture of skills, 49–50, 70
- slavishness, 163n15, 171, 174
- Smith, Adam, 78n23
- sociability
 animality and, 70
 unsociable, 76–79, 77n18, 85n25, 86,
 164–165, 167
- social contract, 194, 198–199
- social evil, 66
- social mores (*Sitten*), 172
- social passions, 80–81
- social sciences, 112
- Sonnenfels, Joseph von, 200n11
- Son of God (ideal of humanity). *see also* grace
 and salvation; Jesus Christ; “well-
 pleasing to God” and other specific
 formulations of the ideal
 change of heart and, 109–110, 115, 126–127,
 130, 138n16
- commentators on, 116n1
- evil propensity and, 72n13
- faith in *reality* of, 126–133, 138n16, 146, 150
- forgiveness and, 135–138
- formulations of the ideal, 116–117
- gender and, 89n1
- grace and salvation and, 148–151
- Jesus Christ and, 148–151
- overview, 115–117
- propensity to evil and, 72n13
- rational and revealed religion and, 19, 149
- symbols and, 117–126, 148–151
- time and, 149
- vicarious atonement and, 144, 145, 146
- Spener, Philip Jacob and his problem, 153–154
- Spinoza, Baruch, 51, 94, 144n3, 191, 197
- state of nature, ethical and political,
 164–166, 193
- state, political
 church and, 194n7, 194–195, 198–199
 coercion of conviction and, 198–199
 coercive enforcement of contracts and, 194
 ethical community and, 168–169
 freedom of conscience and, 206–207
 Judaism as, 192
 just *versus* despotic, 118, 120, 171–172, 182
 Mendelssohn on, 192–195, 194n7, 194–195,
 196, 197, 199–200
 reform and, 183–184
 religious matters and, 196, 197
 universal world state, 166n1
- status, 82
- Stern, Robert, 75n16
- Stoicism, 38, 154
- Stoic maxim, 159
- struggle against evil. *see also* change of heart
 ethical community and, 122, 165–167,
 170–171
- evil propensity and, 104–105
- religion and, 211
- symbolized, 115, 117–120
- students, contemporary, xviii
- subjective grounds, 32–33, 34, 52, 52–56, 67
 “subjectively considered,” 2
- Subjectivity and Irreligion: Atheism and
 Agnosticism* (Ray), xv
- sublimity, 1, 6, 172–174, 173
- submissiveness, 170, 171
- suffering, 119
- Sulzer, J. G., 91
- superiority over others, 78
- supernaturalism. *see also* noumenal world

- anthropomorphism and, 3n3
causality and, 66, 120
Christian doctrine and, 149–151
commentators on, xiv
defined, 18
enthusiasm and, 156
freedom and, 65, 65n3, 129
grace and, 151n7, 162
metaphysical theories and, 64n2
moral freedom and, 157
origin of religion and, 171
religious instruction and, 153
the supersensible, 6, 111, 124, 168. *see also* noumenal world
superstition. *see also* anthropomorphism; miracles
baptism and, 185
church compared, 168
counterfeit service and, 189
defined, 17–18, 120–121
ethical community and, 172
predestination and, 162
religious, 211, 213
revealed faiths and, 170
scripture and, 214
the supersensible and, 168
symbolism *versus*, 124
surplus of merit, 131–132, 133, 134
Sussman, David, 69n8, 97n7
Swinburne, Richard, 29
sycophancy, 122n4
symbols, evil, 7n8
symbols, religious. *see also* aesthetics, Kantian; analogies; anthropomorphism; literalism; possibility; Son of God and other symbols
belief and, 4
change of heart and, 109–110, 125–126
cognition and, 5
doctrines as, 19
ecclesiastical faith and, 17
emotions and, 5, 7n8, 120, 173n4, 177–178
empirical experience and, 4, 4nn4–5, 123–124, 151
faith in, 123
faith in *reality* of Son of God and, 126–127
of God's understanding and will, 28
highest good and, 143
intellectual/emotional appreciation of, xix
interpretation and, 123, 124, 173n4, 213
literalism and, 7n8, 7–8, 123, 173n4, 176, 211, 213
“lost” causes and, 37
need for revealed religion and, 170
Original Sin as, 62
overviews, 4n4, 4–6, 5n6, 7, 117–118, 124–125, 211, 213
perfection and, 15
rational religion and revealed doctrine compared, 17, 120, 126, 140, 149–151
reason and morality and, 214
reform and, 174, 182, 188–189, 211, 214
religion and, 4nn, 4–5, 4–9, 5n6, 9n9, 120–127, 188–189
scripture and tradition and, 6, 7, 188, 201n13
Tillich and, 4n5, 122n4
sympathy, 91, 99, 101, 106
systemization of all teleology, 49–51
- takfir*, 197
talents, 70, 82
taxation for church support, 194n7, 196
technical predisposition, 71
“teleological suspension of the ethical,” 12–13
teleology, systemization of, 49–51
temperament, 100n9, 113n16
temptation, 73, 83–84, 103, 106, 107, 113, 115–116, 149
Tenenbaum, Sergio, 69n8
territories (*Boden, territorium*), 13
thankfulness, 50
theism, 3, 6, 10n10
theistic proofs, 27n1, 27–28
“theistic shrinkage,” 28, 29, 29n3
theocracies, 169
theological faculty at University of Königsberg, 21, 23–24, 25, 25n20
theology, 3, 11, 52, 150
theoretical agnosticism, 59
theoretical cognition, 5, 13n13, 15, 28, 33–34. *see also* possibility
theoretical reasons, 29
things in themselves, 65
thinking for oneself, 25, 25–26, 121, 123, 189. *see also* conscience, freedom of
three Critiques compared, 34–35, 37n10, 45, 51
Tichten und Trachten, 71n12
Tillich, Paul, 4n5, 56, 122n4
time
change of heart and, 89–90, 93–94, 105–106, 110–114, 128n8, 128–129, 130, 131, 146–147
choice of good or evil heart and, 97
dispositions and, 96, 128–130

- time (cont.)
 duty to act from duty and, 128n8
 felicity and, 194
 God's knowledge and, 109
 good heart and, 97
 innateness and, 85
 "kind of person" and, 97
 Original Sin and, 62, 63
 predestination and, 162
 Son of God and, 149
 vicarious atonement and, 146
 "well-pleasing to God" and, 128, 128n8
 timelessness, 128n8
 tobacco industry example, 30n5
 toleration, 11, 14, 191, 212, 213
 tradition, 17, 169–170, 183, 185–187, 201n13,
 202, 212–213. *see also* authority;
 priesthood *and other practices*; scripture;
 service of God
 transcendence, 142
 transcendental idealism, 3n3
 transcendental metaphysics, 64n2
 transcendental theology, 119–120
 transcendent metaphysics, 56, 66, 97, 113,
 116n1, 128n8, 129, 156, 158, 160. *see*
also noumenal world
 translation of Christian doctrines, 123
 treachery, 86
 tribal community, 207, 212
 tribal factors, 35, 190
 Trinity, 149n6, 150
 Troeltsch, Ernst, 188n3
 trust, betrayal of, 70n11
 trust, rational, 52n17
 Truth and Reconciliation in South Africa, 142
 Tutu, Desmond, 142
 2240 C.E., 195, 195n8
 "two perspectives" problem, 128n8
 ultimate end (*letzter Zweck*), 49–51
 unanimity, 167
 unbelief (*Unglaube*), 55
 understanding, 1
 unhappiness, 39, 69
 unity and universality. *see also* "Every man has
 his price."
 church and, 168, 175, 192, 214
 ethical community and, 166, 183n6, 189,
 211–212
 ethical state of nature and, 165
 of evil, 84–86
 Judaism and Christianity and, 203n15
 Mendelssohn and Kant compared, 203
 moral community and, 211–212
 Original Sin and, 62
 propensity to evil and, 85
 reform and, 182, 187, 203, 211,
 212, 213
 of setting highest end, 34
 unity with Christ, 124
 universal credibility, 176
 universal world-state, 166n1
 university, conception of, 23n19, 23–24
 unjust commands, 21, 23
 unreason, 67
 unsocial sociability, 76–79, 77n18, 85n25, 86,
 164, 167
 U.S. Constitution's First Amendment,
 199–200
 U.S. Supreme Court, 205n18
 vanity, 79n23
 Velleman, David, 69n8
 vengeance, 70n11, 81, 176–177
 vicarious satisfaction, 152, 178, 179, 180. *see*
also satisfaction
 vicarious substitute (*Stellvertreter*), 134,
 148, 149
 vices
 of culture, 71
 diabolical, 69, 69n8, 70, 70n11, 79
 inclinations and, 72
 moral and nonmoral incentives and, 108–109
 self-regarding, 79
 theoretical error *versus*, 34
 unsociable sociability and, 77, 77n18
 virtue. *see also* change of heart; moral progress
 acquisition of, 106
 apparent (*virtus phaenomenon*), 92–93, 97,
 105, 106, 107, 108
 degrees of, 41, 73
 faith as, 56n18
 goodness of heart and, 141
 habitual duty and, 92
 inclinations and, 72
 pretense of, 112
 progress and, 104–105
 true (*virtus noumenon*), 92–93, 97, 105,
 106, 108
 visionary (*Phantast*), 42, 43
 vocation (*Bestimmung*), 49–51, 57
 "voluntary (*freiwillig*) determination of our
 judgment," 58
 vulnerable individuals, 35

- wall with Mexico, 7n8
 Walpole's thesis, 75–76, 102–103, 105, 109
 Watkins, Eric, 64n2
 way of thinking (*Denkungsart*), 51, 59, 118,
 119, 120, 124, 125, 126
 weakness of character, 83
 wealth, 171
 Weber, Max, 213
 “well-pleasing to God” (formulation of ideal of
 humanity). *see also* forgiveness; grace
 and salvation
 atonement and, 144, 147
 change of heart and, 17, 88, 94, 109, 115
 counterfeit service and, 187
 faith and, 179–180, 214
 guilt and sin and, 133–134
 hope and, 52, 109, 129–130, 146, 152
 Jesus Christ and, 180
 morality and, 214
 rational and revealed religion compared, 17
 self-acceptance and, 143, 146
 Son of God and, 72n13, 116–117, 125–135,
 146, 149, 150
 as symbol, 137, 150
 time and, 128n8, 128–129
 true service and, 187
*What Does It Mean to Orient Oneself in
 Thinking?* (Kant), 11n11, 34
 “whole end that is natural for every human
 being,” 42, 43
 Wilhelm I of Prussia, 10
 will (free will). *see also* choice; free action;
 “voluntary [*freiwillig*] determination of
 our judgment”
 action and, 31
 “characterization” and, 95
 determination *versus* causality and, 94–96
 diabolical, 68
 God's, 1, 28
 good, 92
 goodness of, 37
 guilt and, 144n3
 highest good and, 47, 54
 holy, 63, 72n13, 116–117, 143
 hope and fear and, 197–198
 moral law and, xx, 1
 noumenal self and, 65, 65n3
 perfect, 127
 sensitive but free and, 65
 vocation and, 50
 wrong, 145–146
 William of Ockham, 69n8
 Williams, Bernard, 65n3
 Willis, Janice Dean, xix
The Will to Believe (James), 58
 Wilmans, C. A., 124n6
 wishful thinking, 32
 wish *versus* choice, 30, 32–33
 witchcraft, 120, 121
Wolfenbüttel Fragments (Reimarus), 10
 Wolff, Christian, 10, 202
 Wöllner, Johann Christoph and his charges, 11,
 20–21, 22–26
 Wolterstorff, Nicholas, 135, 141, 141n2, 147,
 154n9, 180, 180n5, 180–181
 women, 184
 the Word, 149n6
 world-state, 183n6
 worship, acts of, 27n1, 27–28, 122n4
 worthiness of receiving grace. *see* receptivity
 worthiness to be happy
 autonomy and, 40
 ethical community and, 167
 future moral world and, 40
 highest good and, 39
 hope for happiness and, 42, 44, 46,
 46n46
 one faith and, 19
 overview, 37–39
 wrong-doing, 14, 145–146
 xenophobia, *see* tribal factors
 “Your majesty's most loyal subject,” 22
 Zuckert, Rachel, 124n5